

STIC-Biotech/ChemLib

99352

From: Schnizer, Holly
Sent: Tuesday, July 22, 2003 8:17 AM
To: STIC-Biotech/ChemLib
Subject: seq. search for appl. no. 09/444,281

Please search all databases for the following two peptide sequences:

- 1) a peptide having the following sequence from N-terminus SEQ ID NO:35: SEQ ID NO:27: SEQ ID NO:35 ; this peptide is 37 amino acids in length
- 2) a peptide having the following sequence from N-terminus SEQ ID NO:36: SEQ ID NO:27: SEQ ID NO:35; this peptide is 35 amino acids in length

The above is searches for two peptides each made up of three SEQ ID NOs placed in order from N-terminus to C-terminus.

If you have any questions, you may reach me by email or at the phone number listed below.

thank you.

Holly Schnizer
AU 1653
CM1-9E09
305-3722
mailbox: CM1-9B01

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2003, 16:18:05 ; Search time 58.6849 Seconds
(without alignments)
97.370 Million cell updates/sec

Title: SEQ36-27-35

Perfect score: 236

Sequence: 1 ILRWPPWRRKHEAEPAEPILMKKWPWRRK 36

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq_19Jun03:*

- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
- 24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	152.5	64.6	28	21	AAV91800
2	148.5	62.9	27	19	AAW66363
3	130.5	55.3	63	21	AAV44668
4	130.5	55.3	63	21	AAV57142
5	98.5	41.7	21	19	AAV24552
6	98.5	41.7	21	19	AAW24552
7	98.5	41.7	21	19	AAW66376
8	94.5	40.0	15	19	AAV91796
9	94.5	40.0	15	21	AAW66360
					AAV91784
					Amino acid sequenc

10	94	39.8	20	19	AAV24553	Indolicidin analog
11	94	39.8	20	21	AAV91797	Amino acid sequenc
12	93.5	39.6	21	19	AAV24554	Indolicidin analog
13	93.5	39.6	21	19	AAV24554	Amino acid sequenc
14	91	38.6	13	18	AAW12873	Antimicrobial cati
15	91	38.6	13	19	AAV24609	Indolicidin analog
16	91	38.6	13	19	AAW66378	Cationic peptide o
17	91	38.6	13	19	AAW71690	Cationic peptide M
18	91	38.6	13	21	AAV94495	MBI-11 peptide der
19	91	38.6	13	21	AAV92795	Indolicidin analog
20	91	38.6	13	21	AAV91773	Amino acid sequenc
21	91	38.6	13	21	AAV91774	Amino acid sequenc
22	91	38.6	13	21	AAV91818	Amino acid sequenc
23	91	38.6	13	21	AAV91819	Amino acid sequenc
24	91	38.6	13	21	AAV91820	Amino acid sequenc
25	91	38.6	13	23	ABB81254	CP11-NH2 antibacte
26	91	38.6	14	19	AAV24583	Indolicidin analog
27	91	38.6	14	21	AAV91811	Amino acid sequenc
28	91	38.6	21	19	AAV24582	Indolicidin analog
29	91	38.6	21	21	AAV91806	Amino acid sequenc
30	87	36.9	12	19	AAV24580	Indolicidin analog
31	87	36.9	12	21	AAV91804	Amino acid sequenc
32	86	36.4	12	18	AAW12877	Antimicrobial cati
33	86	36.4	12	19	AAV24615	Indolicidin analog
34	86	36.4	12	19	AAV24550	Indolicidin analog
35	86	36.4	12	21	AAV94496	MBI-11B7 peptide d
36	86	36.4	12	21	AAV91791	Amino acid sequenc
37	86	36.4	12	21	AAV91833	Amino acid sequenc
38	86	36.4	13	18	AAW12896	Antimicrobial cati
39	86	36.4	13	19	AAV24613	Indolicidin analog
40	86	36.4	13	19	AAV24572	Indolicidin analog
41	86	36.4	13	21	AAV91803	Amino acid sequenc
42	86	36.4	13	21	AAV91812	Amino acid sequenc
43	86	36.4	14	19	AAV24573	Indolicidin analog
44	86	36.4	14	21	AAV91813	Amino acid sequenc
45	86	36.4	15	18	AAW13802	Antimicrobial cati

ALIGNMENTS

RESULT 1
AAV91800
ID AAV91800 standard; Peptide: 28 AA.
XX
AC AAV91800;
XX
DT 06-JUN-2000 (first entry)
XX
DE Amino acid sequence of cationic peptide MBI 11B20CN.
XX

Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
leukaemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;
breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
multidrug resistance.

Synthetic.

WO9965506-A2.

23-DEC-1999.

14-JUN-1999; 99WO-CA00552.

12-JUN-1998; 98US-0096541.

(MICR-) MICROLOGIX BIOTECH INC.

Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;
WPI; 2000-223549/19.

Novel pharmaceutical composition containing optionally activated

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2003, 16:18:05 ; Search time 58.6849 Seconds
(without alignments)
97.370 Million cell updates/sec

Title: SEQ36-27-35
Perfect score: 236
Sequence: 1 ILRPMWPMRRKHEAPEAPIMLKMPWPMRRK 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: A_Geneseq_19Jun03.*
2: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
23: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
24: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*
25: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	152.5	64.6	28	AA191800	Amino acid sequence
2	148.5	62.9	27	AAW6363	Indolicidin analog
3	130.5	55.3	63	AA144668	Poly-(Indol (1-13)
4	130.5	55.3	63	AA157142	Indolicidin fusion
5	98.5	41.7	21	AA124552	Indolicidin analog
6	98.5	41.7	21	AAW6376	Cationic peptide o
7	98.5	41.7	21	AA191796	Amino acid sequence
8	94.5	40.0	15	AAW6360	Indolicidin analog
9	94.5	40.0	15	AA191784	Amino acid sequence

10	94	39.8	20	AA124553	Indolicidin analog
11	94	39.8	20	AA191797	Amino acid sequence
12	93.5	39.6	21	AA124554	Indolicidin analog
13	93.5	39.6	21	AA191798	Amino acid sequence
14	91	38.6	13	AAW12873	Antimicrobial catl
15	91	38.6	13	AA124609	Indolicidin analog
16	91	38.6	13	AAW6378	Cationic peptide o
17	91	38.6	13	AAW71690	Cationic peptide M
18	91	38.6	13	AA194485	MBI-11 peptide der
19	91	38.6	13	AA192795	Indolicidin analog
20	91	38.6	13	AA191773	Amino acid sequence
21	91	38.6	13	AA191774	Amino acid sequence
22	91	38.6	13	AA191818	Amino acid sequence
23	91	38.6	13	AA191819	Amino acid sequence
24	91	38.6	13	AA191820	Amino acid sequence
25	91	38.6	13	ABW81254	CP11-NH2 antibacte
26	91	38.6	14	AA124583	Indolicidin analog
27	91	38.6	14	AA191811	Amino acid sequence
28	91	38.6	21	AA124582	Indolicidin analog
29	91	38.6	21	AA191806	Amino acid sequence
30	87	36.9	12	AA124580	Indolicidin analog
31	87	36.9	12	AA191804	Amino acid sequence
32	86	36.4	12	AAW12877	Antimicrobial catl
33	86	36.4	12	AA124615	Indolicidin analog
34	86	36.4	12	AA124550	Indolicidin analog
35	86	36.4	12	AA194496	MBI-11B7 peptide d
36	86	36.4	12	AA191791	Amino acid sequence
37	86	36.4	12	AA191833	Amino acid sequence
38	86	36.4	13	AAW12896	Antimicrobial catl
39	86	36.4	13	AA124613	Indolicidin analog
40	86	36.4	13	AA124572	Indolicidin analog
41	86	36.4	13	AA191803	Amino acid sequence
42	86	36.4	13	AA191812	Amino acid sequence
43	86	36.4	14	AA124573	Indolicidin analog
44	86	36.4	14	AA191813	Amino acid sequence
45	86	36.4	15	AAW13802	Antimicrobial catl

ALIGNMENTS

RESULT 1
AA191800
ID AA191800 standard; Peptide: 28 AA.
XX
AC AA191800;
XX
DT 06-JUN-2000 (first entry)
XX
DE Amino acid sequence of cationic peptide MBI_11B20CN.
XX
KW Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
KW Leukaemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;
KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
KW multidrug resistance.
XX
OS Synthetic.
XX
FN W09965506-A2.
XX
PD 23-DEC-1999.
XX
PF 14-JUN-1999; 99WO-CA00552.
XX
PR 12-JUN-1998; 98US-0096541.
XX
PA (MICR-) MICROLOCIX BIOTECH INC.
XX
PI Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;
XX WPI: 2000-223549/19.
XX
PT Novel pharmaceutical composition containing optionally activated

PT polyoxalkylene-modified cationic peptides, useful for treating tumours
 XX
 PS Claim 1; Page 15; 94pp; English.
 CC
 CC This sequence represents a cationic peptide amino acid sequence, which
 CC can be used in the pharmaceutical composition of the invention. The
 CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxalkylene (APO)-modified cationic peptide. The
 CC modification of peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon.
 CC
 SQ Sequence 28 AA:
 Query Match 64.6%; Score 152.5; DB 21; Length 28;
 Best Local Similarity 63.9%; Pred. No. 1.6e-12;
 Matches 23; Conservative 1; Mismatches 1; Indels 11; Gaps 1;
 Seq ID No: 36
 1 ILRMPMPWRKHEAPEPIMILKKMPMPWRK,36
 1 ILRMPMPWRK-----ILRMPMPWRK 25
 Db
 RESULT 2
 AAM66363
 ID AAM66363 standard; peptide; 27 AA.
 XX
 AC AAM66363;
 XX
 DT 12-JAN-1999 (first entry)
 XX
 DE Indolicidin analogue MBI 11B20.
 XX
 DE Indolicidin analogue; resistance; cationic peptide; antibiotic;
 KM bacterial infection; tolerance; antibacterial; microorganism;
 KM bacteria; fungus; parasite; virus.
 XX
 OS Bos taurus.
 OS Synthetic.
 OS
 XX WO9840401-A2.
 XX
 PD 17-SEP-1998.
 XX
 PF 10-MAR-1998; 98WO-CA00190.
 XX
 PR 25-FEB-1998; 98US-0030619.
 PR 10-MAR-1997; 97US-0040649.
 PR 20-AUG-1997; 97US-0915314.
 PR 26-SEP-1997; 97US-0060099.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.
 XX
 PI Fraser JR, McNICOL PJ, West MHP;
 XX
 DR WPI; 1998-520800/44.
 XX
 PT New indolicidin peptide analogues - useful for, e.g. enhancing
 PT activity of antibiotic or overcoming tolerance, acquired resistance
 PT or inherent resistance of microorganisms
 XX
 PS Claim 1; Page 91; 105pp; English.
 XX
 CC The present sequence represents an indolicidin analogue. The present
 CC invention describes compositions and methods for treating infection,
 CC especially bacterial infections. The compositions and methods use
 CC cationic peptides in combination with an antibiotic agent which are
 CC then administered to a patient to enhance the activity of the antibiotic
 CC agent, to overcome: (a) tolerance; (b) acquired resistance; and (c)
 CC inherent resistance. The combinations of antibiotics and cationic

CC peptides can provide synergistic activity against a microorganism that
 CC is tolerant, inherently resistant, or has acquired resistance to an
 CC antibiotic agent. They can be used for killing e.g. bacteria, fungi,
 CC parasites and viruses.
 CC
 SQ Sequence 27 AA:
 Query Match 62.9%; Score 148.5; DB 19; Length 27;
 Best Local Similarity 66.7%; Pred. No. 4.9e-12;
 Matches 24; Conservative 1; Mismatches 0; Indels 11; Gaps 2;
 Seq ID No: 36
 1 ILRMPMPWRKHEAPEPIMILKKMPMPWRK 36
 1 ILRMPMPWRK-----ILRMPMPWRK 25
 Db
 RESULT 3
 AAY44668
 ID AAY44668 standard; Protein; 63 AA.
 XX
 AC AAY44668;
 XX
 DT 18-APR-2000 (first entry)
 XX
 DE Poly-(Indol (1-13)-Met-Ala-Arg-Ile-Ala-Met)3 protein.
 XX
 XX Crosslinked indolicidin analog; X-indolicidin; poly-Indol 1-13;
 KM stability; bovine neutrophil; antimicrobial; antibacterial; fungicide;
 KM protozoacide; virucide; anti-HIV; human immunodeficiency virus-1;
 KM HIV-1; gram positive bacteria; gram negative; Staphylococcus aureus;
 KM Escherichia coli; Salmonella typhimurium; yeast; fungi; protozoa;
 KM Candida albicans; Cryptococcus neoformans; Giardia; Acanthamoeba;
 KM hexapeptide spacer.
 XX
 OS Synthetic.
 OS Bos sp.
 XX
 FH Key
 FT Region
 FT Location/Qualifiers
 FT 1..5 Enterokinase_recognition_site
 FT /label= Enterokinase_recognition_site
 FT 5..6
 FT /label= Enterokinase_cleavage_site
 FT 6..7
 FT /label= Cyanogen_bromide_cleavage_site
 FT 20..21
 FT /label= Cyanogen_bromide_cleavage_site
 FT 25..26
 FT /label= Cyanogen_bromide_cleavage_site
 FT 39..40
 FT /label= Cyanogen_bromide_cleavage_site
 FT 44..45
 FT /label= Cyanogen_bromide_cleavage_site
 FT 58..59
 FT /label= Cyanogen_bromide_cleavage_site
 FT 20..25
 FT /label= Hexapeptide_spacer
 FT 39..44
 FT /label= Hexapeptide_spacer
 FT 58..63
 FT /label= Hexapeptide_spacer
 FT Region
 XX
 PN WO965510-A1.
 XX
 PD 23-DEC-1999.
 XX
 PF 20-MAY-1999; 99WO-US11165.
 XX
 PR 18-JUN-1998; 98US-0093631.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Selsted ME, Osapay K;
 XX

DR WPI: 2000-147133/13.
DR N-PSDB; AAZ49764.
XX
PT Crosslinked indolicidin analogs with antimicrobial activity against
PT bacteria, yeast, fungi, protozoa and viruses
XX
PS Example 1C; Fig 1; 53pp; English.
XX
CC The patent discloses crosslinked analogs of indolicidin (Indol 1-13)
CC which is a naturally occurring peptide isolated from bovine neutrophils
CC and has antimicrobial activity. The crosslinked indolicidin
CC (X-indolicidin) analogs are stable and have antimicrobial activity
CC against gram positive and negative bacteria (e.g. Staphylococcus aureus,
CC Escherichia coli and Salmonella typhimurium), yeasts and fungi (e.g.
CC Candida albicans, Cryptococcus neoformans), protozoa (e.g. Giardia
CC species and Acanthamoeba species), and viruses (e.g. HIV-1).
CC They can be used for reducing or inhibiting the growth or survival of
CC microorganisms in an environment e.g. a food or food product, a
CC solution, an inanimate object comprising a surface, or a mammal.
CC The present sequence is a protein comprising three
CC copies of Indol 1-13 each separated by a hexapeptide spacer sequence.
CC A recombinant construct encoding this sequence was used for the
CC expression of indol-homoserine (Hse) analog. The ability of
CC indol-Hse analog to maintain antimicrobial activity provides a means to
CC produce X-indolicidin analog precursors in sufficient quantities.
XX
SO Sequence 63 AA:
XX
Query Match 55.3%; Score 130.5; DB 21; Length 63;
Best Local Similarity 61.8%; Pred. No. 2.4e-09;
Matches 21; Conservative 1; Mismatches 5; Indels 7; Gaps 2;
XX
OY 3 RWPMPWRRKHAEPEAPIMILK-KWPMPWRR 35
DB 11 KWPMPWRRM-----ARIAMILPKWMPWRR 38
XX
RESULT 4
AAV57142
ID AAV57142 standard; Protein: 63 AA.
XX
AC AAV57142;
XX
DT 28-FEB-2000 (first entry)
XX
DE Indolicidin fusion peptide amino acid sequence.
XX
XX Indolicidin analogue; antimicrobial activity; helminth; bacteria; virus;
KW treatment; inhibit growth; micro-organism; contact lens solution;
KW transgenic plant; surgical instrument; yeast; fungi; protozoa.
XX
OS Synthetic.
XX
PN WO9958141-A1.
XX
PD 18-NOV-1999.
XX
PE 05-MAY-1999; 99WO-US09942.
XX
PR 12-MAY-1998; 98US-0076227.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Sealed ME;
XX
DR WPI: 2000-053028/04.
DR N-PSDB; AAZ45123.
XX
PT New indolicidin analogues, active against bacteria, yeast, fungi,
PT protozoa and virus, used for, e.g. treating infections
XX
PS Disclosure: Fig 6; 62pp; English.
XX

CC This is the amino acid sequence of an example of a fusion protein which
CC consists of an indolicidin analogue linked to another peptide.
CC Peptides AAV57109-Y57138 and AAV57143-Y57144 are new indolicidin
CC analogues, which have a homoserine residue and/or a truncated amino
CC terminal region. The analogues have the following amino acid sequence:
CC Xaa1-Xaa2-Xaa3-Xaa5-Xaa6-Pro-Xaa6-Pro-Xaa6-Xaa7-Xaa8
CC Where:
CC Xaa1 = Ile, Leu, Val, Ala, Gly or absent;
CC Xaa2 = Ile, Leu, Val, Ala, Gly or absent;
CC Xaa3 = Pro or absent;
CC Xaa4 = Trp, Phe or absent;
CC Xaa5 = Arg, Lys or absent;
CC Xaa6 = Trp or Phe;
CC Xaa7 = Arg, Lys or absent;
CC Xaa8 = homoserine (Hse), Met, Met-Xaa9-Met or absent, and
CC Xaa9 = at least one amino acid;
CC provided that if Xaa1 is present, Xaa8 = Hse, Met or Met-Xaa9-Met;
CC and further provided that: if Xaa2 is absent, Xaa1 is absent; if Xaa3 is
CC absent, Xaa1 and Xaa2 are absent; if Xaa4 is absent, Xaa1, Xaa2 and Xaa3
CC are absent; and if Xaa5 is absent, Xaa1, Xaa2, Xaa3 and Xaa4 are absent.
CC The indolicidin analogues can be used to create a fusion polypeptide
CC consisting of the analogue linked to a peptide. The indolicidin
CC analogues have antimicrobial activity against gram positive bacteria,
CC gram negative bacteria, yeast, fungus, protozoa and viruses (e.g. HIV-1).
CC They are also active against helminths. The analogues can be used for
CC reducing or inhibiting growth or survival of a microorganism. They can be
CC used for treating infections. They can also be included in a liquid such
CC as water or an aqueous solution, e.g. contact lens solution. The
CC analogues have potential uses in food products, and in objects such as
CC the surface of an instrument used to prepare food or to perform surgery.
CC Transgenic plants or animals useful in the food industry can be produced
CC by introducing a nucleic acid molecule encoding an indolicidin analogue
CC into the germ-line cells of such organisms.
XX
SO Sequence 63 AA:
XX
Query Match 55.3%; Score 130.5; DB 21; Length 63;
Best Local Similarity 61.8%; Pred. No. 2.4e-09;
Matches 21; Conservative 1; Mismatches 5; Indels 7; Gaps 2;
XX
OY 3 RWPMPWRRKHAEPEAPIMILK-KWPMPWRR 35
DB 11 KWPMPWRRM-----ARIAMILPKWMPWRR 38
XX
RESULT 5
AAV24552
ID AAV24552 standard; peptide: 21 AA.
XX
AC AAV24552;
XX
DT 18-AUG-1999 (first entry)
XX
DE Indolicidin analogue #4.
XX
XX Indolicidin; bacterial infection; photo-oxidised solubiliser;
KW antimicrobial; antibiotic; antiarrhythmic; surface disinfectant;
KW additive; shampoo; soap; insecticide; herbicide; preservative;
KW food; technical material.
XX
OS Synthetic.
XX
PN WO9807745-A2.
XX
PD 26-FEB-1998.
XX
PE 21-AUG-1997; 97WO-US14779.
XX
PR 13-JAN-1997; 97US-0034949.
PR 21-AUG-1996; 96US-0024754.
XX
PA (MICR-) MICROLOGIX BIOTECH INC.
XX

PI Effle D, Fraser JR, Krieger TJ, Taylor R, West MH;
 XX WPI; 1998-169090/15.
 XX
 PT New indolicidin analogues with antimicrobial activity and related
 PT nucleic acid - vectors, transformed cells and antibodies, also
 PT conjugates with polyoxyalkylene glycol and fatty acid to reduce
 XX toxicity, useful therapeutically, as disinfectants etc.
 PS Claim 11; Page 88; 129pp; English.
 XX
 CC AAY24549 to AAY24615 represent indolicidin analogues of formulae
 CC (I)-(VII) containing up to 25 amino acids (aa): R₁XXXXXB (I), B₂XXXXXB
 CC (II), B₃XXXXXB (III), B₄XXXXXB (IV), B₅XXXXXB (V), B₆XXXXXB (VI),
 CC (V), LBH₁XXXXXB (VII), LK₁XXXXXB (VIII) and B₇XXXXXB (IX).
 CC Where Z = P or V; X = hydrophobic residue, preferably W; B = basic aa,
 CC preferably R or K; AA = any aa; n = 0 or 1; in (II), at least 1 Z = V;
 CC in (VIII) at least 2 X = F or Y. The analogues are used to treat
 CC infections caused by bacteria (Gram positive or negative, or anaerobic);
 CC fungi (yeast or moulds); parasites (protozoa, nematodes, cestodes or
 CC trematodes) or viruses. Typical of very many pathogens that can be
 CC controlled are Leishmania, Trypanosoma, Ascaris lumbricoides, Fasciola
 CC hepatica, Klebsiella pneumoniae, Bordetella pertussis, Staphylococcus
 CC aureus, Listeria, Clostridium, rotavirus and papilloma virus. Compounds
 CC derived from the analogues may be used similarly; the compounds may
 CC also be prepared from antibiotics or antiparasitic agents. The analogues
 CC may be used therapeutically or to coat medical devices; also they are
 CC useful as surface disinfectants, as additives to shampoo or soaps, as
 CC insecticides or herbicides, or as preservatives for foods and technical
 CC materials. The analogues are administered by injection, lavage, orally
 CC or topically, generally at 0.1-50 mg/Kg. These analogues have a broader
 CC spectrum of activity than indolicidin and modification as compounds
 CC reduces their toxicity.
 CC
 SQ Sequence. 21 AA;
 Query Match 41.7%; Score 98.5; DB 19; Length 21;
 Best Local Similarity 66.7%; Pred. No. 7.7e-06;
 Matches 18; Conservative 0; Mismatches 0; Indels 9; Gaps 1;
 QY 1 ILRMPMPWRRKHEAPEAPIMILKK 27
 Db 1 ILRMPMPWRRKHEAPEAPIMILKK 18
 RESULT 6
 AAM6376
 ID AAM6376 standard; peptide; 21 AA.
 AC AAM6376;
 XX
 DT 12-JAN-1999 (first entry)
 DE Cationic peptide of claim 15 #3.
 XX
 KW Indolicidin analogue; resistance; cationic peptide; antibiotic;
 KW bacterial infection; tolerance; antibacterial; microorganism;
 KW bacteria; fungus; parasite; virus.
 XX
 OS Synthetic.
 XX
 PN WO9840401-A2.
 PD 17-SEP-1998.
 XX
 PF 10-MAR-1998; 98WO-CA00190.
 XX
 PR 25-FEB-1998; 98US-0030619.
 PR 10-MAR-1997; 97US-0040649.
 PR 20-AUG-1997; 97US-0915314.
 PR 26-SEP-1997; 97US-0060099.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.

XX
 PI Fraser JR, McNicol PJ, West MHP;
 XX WPI; 1998-520800/44.
 XX
 PT New indolicidin peptide analogues - useful for, e.g. enhancing
 PT activity of antibiotic or overcoming tolerance, acquired resistance
 PT or inherent resistance of microorganisms
 XX
 PS Claim 15; Page 93; 105pp; English.
 XX
 CC The present sequence represents a specifically claimed cationic peptide
 CC from the present invention. The present invention describes compositions
 CC and methods for treating infection, especially bacterial infections. The
 CC compositions and methods use cationic peptides in combination with an
 CC antibiotic agent which are then administered to a patient to enhance the
 CC activity of the antibiotic agent, to overcome: (a) tolerance; (b)
 CC acquired resistance; and (c) inherent resistance. The combinations of
 CC antibiotics and cationic peptides can provide synergistic activity
 CC against a microorganism that is tolerant, inherently resistant, or has
 CC acquired resistance to an antibiotic agent. They can be used for killing
 CC e.g. bacteria, fungi, parasites and viruses.
 CC
 SQ Sequence 21 AA;
 Query Match 41.7%; Score 98.5; DB 19; Length 21;
 Best Local Similarity 66.7%; Pred. No. 7.7e-06;
 Matches 18; Conservative 0; Mismatches 0; Indels 9; Gaps 1;
 QY 1 ILRMPMPWRRKHEAPEAPIMILKK 27
 Db 1 ILRMPMPWRRKHEAPEAPIMILKK 18
 RESULT 7
 AAY91796
 ID AAY91796 standard; Peptide; 21 AA.
 AC AAY91796;
 XX
 DT 06-JUN-2000 (first entry)
 DE Amino acid sequence of cationic peptide MBI 11B16CN.
 XX
 KW Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
 KW leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma;
 KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
 KW multidrug resistance.
 XX
 OS Synthetic.
 XX
 PN WO9965506-A2.
 PD 23-DEC-1999.
 XX
 PF 14-JUN-1999; 99WO-CA00552.
 PF 12-JUN-1998; 98US-0096541.
 PR
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.
 XX
 PI Friedland HD, Krieger TJ, Taylor R, Effle D, Fraser JR, West MHP;
 XX WPI; 2000-223549/19.
 XX
 PT Novel pharmaceutical composition containing optionally activated
 PT polyoxyalkylene-modified cationic peptides, useful for treating tumours
 PS
 XX Disclosure; Page 15; 94pp; English.
 CC This sequence represents a cationic peptide amino acid sequence, which
 CC can be used in the pharmaceutical composition of the invention. The

CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxalkylene (APO)-modified cationic peptide. The
 CC modification of peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon.

XX Sequence 21 AA;

Query Match 41.7%; Score 98.5; DB 21; Length 21;
 Best Local Similarity 66.7%; Pred. No. 7.7e-06;
 Matches 18; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

OY 1 ILRPMWPMRRKHEAPEAPIMILKK 27
 |||||
 DB 1 ILRPMWPMRRK-----IMILKK 18

RESULT 8
 AAW66360
 ID AAW66360 standard; peptide; 15 AA.

XX AAW66360;

DT 12-JAN-1999 (first entry)

XX Indolicidin analogue MBI 11A9.

DE Indolicidin analogue; resistance; cationic peptide; antibiotic;
 KM bacterial infection; tolerance; antibacterial; microorganism;
 KW bacteria; fungus; parasite; virus.

XX Bos taurus.

OS Synthetic.

XX WO9840401-A2.

XX 17-SEP-1998.

XX 10-MAR-1998; 98WO-CA00190.

XX 25-FEB-1998; 98US-0030619.

XX 10-MAR-1997; 97US-0040649.

XX 20-AUG-1997; 97US-0915314.

XX 26-SEP-1997; 97US-0060099.

XX (MICR-) MICROLOGIX BIOTECH INC.

XX Fraser JR, McNicol PJ, West MHP;

XX WPI: 1998-520800/44.

XX New indolicidin peptide analogues - useful for, e.g. enhancing
 PT activity of antibiotic or overcoming tolerance, acquired resistance
 PT or inherent resistance of microorganisms

XX Claim 1; Page 91; 105pp; English.

CC The present sequence represents an indolicidin analogue. The present
 CC invention describes compositions and methods for treating infection,
 CC especially bacterial infections. The compositions and methods use
 CC cationic peptides in combination with an antibiotic agent which are
 CC then administered to a patient to enhance the activity of the antibiotic
 CC agent, to overcome: (a) tolerance; (b) acquired resistance; and (c)
 CC inherent resistance. The combinations of antibiotics and cationic
 CC peptides can provide synergistic activity against a microorganism that
 CC is tolerant, inherently resistant, or has acquired resistance to an
 CC antibiotic agent. They can be used for killing e.g. bacteria, fungi,
 CC parasites and viruses.

XX Sequence 15 AA;

Query Match 40.0%; Score 94.5; DB 19; Length 15;
 Best Local Similarity 41.7%; Pred. No. 1.7e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 21; Gaps 1;

OY 1 ILRPMWPMRRKHEAPEAPIMILKKPMWPMRRK 36
 |||||
 DB 1 ILRPMW-----WPMWPMRRK 15

RESULT 9
 AAY91784

ID AAY91784 standard; peptide; 15 AA.

XX AAY91784;

DT 06-JUN-2000 (first entry)

XX Amino acid sequence of cationic peptide MBI 11A9CN.

XX Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
 KM leukaemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;
 KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
 KM multidrug resistance.

XX Synthetic.

XX WO9965506-A2.

XX 23-DEC-1999.

XX 14-JUN-1999; 99WO-CA00552.

XX 12-JUN-1998; 98US-0096541.

XX (MICR-) MICROLOGIX BIOTECH INC.

XX Friedland HD, Krieger TJ, Taylor R, Effle D, Fraser JR, West MHP;

XX WPI: 2000-223549/19.

XX Novel pharmaceutical composition containing optionally activated
 PT polyoxalkylene-modified cationic peptides, useful for treating tumours

XX Claim 1; Page 14; 94pp; English.

CC This sequence represents a cationic peptide amino acid sequence, which
 CC can be used in the pharmaceutical composition of the invention. The
 CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxalkylene (APO)-modified cationic peptide. The
 CC modification of peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon.

XX Sequence 15 AA;

Query Match 40.0%; Score 94.5; DB 21; Length 15;
 Best Local Similarity 41.7%; Pred. No. 1.7e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 21; Gaps 1;

OY 1 ILRPMWPMRRKHEAPEAPIMILKKPMWPMRRK 36
 |||||
 DB 1 ILRPMW-----WPMWPMRRK 15

RESULT 10
 AAY24553

ID AAY24553 standard; peptide; 20 AA.

XX AAY24553;

DT 18-AUG-1999 (first entry)
 XX Indolicidin analogue #5.
 XX
 XX Indolicidin: bacterial infection; photo-oxidised solubiliser;
 KW antimicrobial; antibiotic; antitarrhythmic; surface disinfectant;
 KW additive; shampoo; soap; insecticide; herbicide; preservative;
 KW food; technical material.
 OS Synthetic.
 XX
 XX WO9807745-A2.
 XX
 XX 26-FEB-1998.
 XX
 XX 21-AUG-1997; 97MO-US14779.
 XX
 XX 13-JAN-1997; 97US-0034949.
 XX 21-AUG-1996; 96US-0024754.
 XX
 XX (MICR-) MICROLOGIX BIOTECH INC.
 XX
 XX Erfile D, Fraser JR, Krieger TJ, Taylor R, West MH;
 XX WPI; 1998-169090/15.
 XX
 XX New indolicidin analogues with antimicrobial activity and related
 PT nucleic acid - vectors, transformed cells and antibodies, also
 PT conjugates with polyoxalkylene glycol and fatty acids to reduce
 PT toxicity, useful therapeutically, as disinfectants etc.
 XX
 PS Claim 11: Page 86; 129pp; English.
 XX
 CC AAY24549 to AAY24615 represent indolicidin analogues of formulae
 CC (I)-(VII) containing up to 25 amino acids (aa): RXXXXXB (I), BXXXXXB
 CC (II), BBBBBXXXXB (III), BXXXXBBBn(AA)nMIBBGS (IV), BXXXXXB(AA)nM
 CC (V), LBnXXXXnXXXXX (VI), LKXXXXXRRK (VII) and BXXXXXBnBB (VIII).
 CC Where 2 = P or V; X = hydrophobic residue, preferably W; B = basic aa;
 CC preferably R or K; AA = any aa; n = 0 or 1; in (II), at least 1 Z = V;
 CC in (VIII) at least 2 X = F or Y. The analogues are used to treat
 CC infections caused by bacteria (Gram positive or negative, or anaerobic);
 CC fungi (yeast or moulds); parasites (protozoa, nematodes, cestodes or
 CC trematodes) or viruses. Typical of very many pathogens that can be
 CC controlled are Leishmania, Trypanosoma, Ascaris lumbricoides, Fasciola
 CC hepatica, Klebsiella pneumoniae, Bordetella pertussis, Staphylococcus
 CC aureus, Listeria, Clostridium, rotavirus and papilloma virus. Compounds
 CC derived from the analogues may be used similarly; the compounds may
 CC also be prepared from antibiotics or antitarrhythmic agents. The analogues
 CC may be used therapeutically or to coat medical devices; also they are
 CC useful as surface disinfectants, as additives to shampoo or soaps, as
 CC insecticides or herbicides, or as preservatives for foods and technical
 CC materials. The analogues are administered by injection, lavage, orally
 CC or topically, generally at 0.1-50 mg/Kg. These analogues have a broader
 CC spectrum of activity than indolicidin and modification as compounds
 CC reduces their toxicity.
 CC
 SQ Sequence 20 AA:
 Query Match 39.8%; Score 94; DB 19; Length 20;
 Best Local Similarity 63.0%; Pred. No. 2.7e-05;
 Matches 17; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
 OY 1 ILRPMWMPWRRKHEAPEEPIMLKK 27
 DB 1 ILRPMWMPWRRK-----MILKK 17
 RESULT 11
 AAY91797
 ID AAY91797 standard; Peptide: 20 AA.
 XX
 AC AAY91797;
 XX

DT 06-JUN-2000 (first entry)
 XX Amino acid sequence of cationic peptide MBI 11B17CN.
 DE
 XX
 XX Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
 KW leukaemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;
 KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
 KW multidrug resistance.
 XX
 OS Synthetic.
 XX
 XX WO965506-A2.
 XX
 XX 23-DEC-1999.
 XX
 XX 14-JUN-1999; 99MO-CA00552.
 XX
 XX 12-JUN-1998; 98US-0096541.
 XX
 XX (MICR-) MICROLOGIX BIOTECH INC.
 XX
 XX Friedland HD, Krieger TJ, Taylor R, Erfile D, Fraser JR, West MHP;
 XX WPI; 2000-223549/19.
 XX
 XX Novel pharmaceutical composition containing optionally activated
 PT polyoxalkylene-modified cationic peptides, useful for treating tumours
 PT
 PS Disclosure; Page 15; 94pp; English.
 XX
 CC This sequence represents a cationic peptide amino acid sequence, which
 CC can be used in the pharmaceutical composition of the invention. The
 CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxalkylene (APO)-modified cationic peptide. The
 CC modification of peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon.
 CC
 SQ Sequence 20 AA:
 Query Match 39.8%; Score 94; DB 21; Length 20;
 Best Local Similarity 63.0%; Pred. No. 2.7e-05;
 Matches 17; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
 OY 1 ILRPMWMPWRRKHEAPEEPIMLKK 27
 DB 1 ILRPMWMPWRRK-----MILKK 17
 RESULT 12
 AAY24554
 ID AAY24554 standard; peptide: 21 AA.
 XX
 AC AAY24554;
 XX
 XX 18-AUG-1999 (first entry)
 DE
 XX Indolicidin analogue #6.
 XX
 XX Indolicidin: bacterial infection; photo-oxidised solubiliser;
 KW antimicrobial; antibiotic; antitarrhythmic; surface disinfectant;
 KW additive; shampoo; soap; insecticide; herbicide; preservative;
 KW food; technical material.
 OS Synthetic.
 XX
 XX WO9807745-A2.
 XX
 XX 26-FEB-1998.
 XX

PF 21-AUG-1997: 97WO-US14779.
 XX
 PR 13-JAN-1997: 97US-0034949.
 PR 21-AUG-1996: 96US-0024754.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.
 PI Erfle D, Fraser JR, Krieger TJ, Taylor R, West MH;
 XX WPI: 1998-169090/15.
 DR
 XX
 PT New indolicidin analogues with antimicrobial activity and related
 PT nucleic acid - vectors, transformed cells and antibodies, also
 PT conjugates with polyoxalkylene glycol and fatty acid to reduce
 PT toxicity, useful therapeutically, as disinfectants etc.
 XX
 PS Claim 11; Page 88; 129pp: English.
 CC
 CC AAY24549 to AAY24615 represent indolicidin analogues of formulae
 CC (I)-(VIII) containing up to 25 amino acids (aa): RXXXZXB (I), BXXXXXB
 CC (II), BBRXXXXXB (III), BXXXXBBB(AA)MILBAGS (IV), BXXXXXB(AA)nm
 CC (V), LBnXXXXXnX (VI), LKXXXXXnX (VII) and BXXXXXB (VIII).
 CC Where Z = F or V; X = hydrophobic residue, preferably W; B = basic aa,
 CC preferably R or K; AA = any aa; n = 0 or 1; in (II), at least 1 Z = V;
 CC in (VIII) at least 2 X = F or Y. The analogues are used to treat
 CC infections caused by bacteria (Gram positive or negative, or anaerobic);
 CC fungi (yeast or moulds); parasites (protozoa, nematodes, cestodes or
 CC trematodes) or viruses. Typical of very many pathogens that can be
 CC controlled are Leishmania, Trypanosoma, Ascaris lumbricoides, Fasciola
 CC hepatica, Klebsiella pneumoniae, Bordetella pertussis, Staphylococcus
 CC aureus, Listeria, Clostridium, rotavirus and papilloma virus. Compounds
 CC derived from the analogues may be used similarly; the compounds may
 CC also be prepared from antibiotics or antiarrhythmic agents. The analogues
 CC may be used therapeutically or to coat medical devices; also they are
 CC useful as surface disinfectants, as additives to shampoo or soaps, as
 CC insecticides or herbicides, or as preservatives for foods and technical
 CC materials. The analogues are administered by injection, lavage, orally
 CC or topically, generally at 0.1-50 mg/kg. These analogues have a broader
 CC spectrum of activity than indolicidin and modification as compounds
 CC reduces their toxicity.
 CC
 SO Sequence 21 AA:
 Query Match 39.6%; Score 93.5; DB 19; Length 21;
 Best Local Similarity 63.0%; Pred. No. 3.3e-05;
 Matches 17; Conservative 0; Mismatches 1; Indels 9; Gaps 1;
 QY 1 ILRPMWPMWRKHEAPEAEPIMILKK 27
 1 ILRPMWPMWRKHEAPEAEPIMILKK 18
 DB 1 ILRPMWPMWRKHEAPEAEPIMILKK 18
 RESULT 13
 AAY91798
 ID AAY91798 standard; Peptide; 21 AA.
 XX
 AC AAY91798;
 XX
 DT 06-JUN-2000 (first entry)
 XX
 XX Amino acid sequence of cationic peptide MBI 11B18CN.
 DE
 XX
 KW Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
 KW leukaemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;
 KW breast; lung; ovary; uterus; skin; prostate; liver; colon;
 KW multidrug resistance.
 XX
 OS Synthetic.
 XX
 XX WO9965506-A2.
 PN
 PD 23-DEC-1999.
 XX

PF 14-JUN-1999: 99WO-CA00552.
 XX
 PR 12-JUN-1998: 98US-0096541.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.
 PI Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;
 XX WPI: 2000-223549/19.
 DR
 XX
 PT Novel pharmaceutical composition containing optionally activated
 PT polyoxalkylene-modified cationic peptides, useful for treating tumours
 XX
 PS Disclosure: Page 15; 94pp: English.
 CC
 CC This sequence represents a cationic peptide amino acid sequence, which
 CC can be used in the pharmaceutical composition of the invention. The
 CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxalkylene (APO)-modified cationic peptide. The
 CC modification of peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon.
 CC
 SO Sequence 21 AA:
 Query Match 39.6%; Score 93.5; DB 21; Length 21;
 Best Local Similarity 63.0%; Pred. No. 3.3e-05;
 Matches 17; Conservative 0; Mismatches 1; Indels 9; Gaps 1;
 QY 1 ILRPMWPMWRKHEAPEAEPIMILKK 27
 1 ILRPMWPMWRKHEAPEAEPIMILKK 18
 DB 1 ILRPMWPMWRKHEAPEAEPIMILKK 18
 RESULT 14
 AAM12873
 ID AAM12873 standard; peptide; 13 AA.
 XX
 AC AAM12873;
 XX
 DT 10-DEC-1997 (first entry)
 XX
 DE Antimicrobial cationic peptide CP-11.
 XX
 KW Bacterial; viral; antitumour; food; preservative; inhibitor; growth;
 KW bacterium; yeast; endotoxaemia; sepsis; antibiotic; fungal;
 KW antiviral; Candida albicans; steriliant; Salmonella; Yersinia;
 KW Shigella.
 XX
 OS Synthetic.
 XX
 XX WO9708199-A2.
 PN
 PD 06-MAR-1997.
 XX
 PF 23-AUG-1996: 96WO-IB00996.
 PR
 PR 23-AUG-1995: 95US-0002687.
 XX
 PA (UYBR-) UNIV BRITISH COLUMBIA.
 PI Falla TJ, Gough M, Hancock RW;
 XX WPI: 1997-179179/16.
 DR
 XX
 PT Cationic peptide(s) having anti-microbial activity - used for the
 PT inhibition of bacterial and viral growth, as an antitumour agent,
 PT and as a food preservative
 XX
 PS Claim 2; Page 65; 89pp: English.

XX The present sequence represents a specifically claimed novel isolated
 CC cationic peptide which has antimicrobial activity. The amino acid
 CC sequence of antimicrobial cationic peptides (including the present
 CC sequence) is selected from: $X_1X_1Prox_2X_3X_2Pro(x_2X_2Pro)_X_2X_3(X_3)_O$;
 CC $X_1X_1P(x_2X_3X_4)(X_5)P(x_2X_3X_3X_3); X_1X_1X_3(ProTrp)x_3X_2X_5X_2X_5X_2(X_5)_O$;
 CC $X_1X_1X_3X_3X_2Pro(x_2X_2Pro)(nX_2)(X_5)m$; where $m = 1-5$; $n = 1-2$; $o = 2-5$; r
 CC $= 0-8$; $u = 0-1$; $X_1 = Ile, Leu, Val, Phe, Tyr, Trp$ or Met ; $X_2 = Trp$ or
 CC Pro ; $X_3 = Arg$ or Lys ; $X_4 = Trp$ or Lys ; and $X_5 = Phe, Trp, Arg, Lys$ or
 CC Pro . The peptides are preferably amidated or carboxymethylated. The
 CC peptides may be used in methods for inhibiting the growth of a bacterium
 CC or yeast, or for inhibiting an endotoxaemia or sepsis associated
 CC disorder in a subject. The peptides have a broad activity against
 CC antibiotic resistant bacteria, combined with activity against the
 CC medically important fungus *Candida albicans*. In addition, the peptides
 CC are useful as antitumour agents and/or antiviral agents. The peptides
 CC may be used as sterilants or preservatives of materials susceptible to
 CC microbial or viral contamination, e.g. in processed foods to inhibit
 CC *Salmonella*, *Yersinia* and *Shigella*. The peptides are compact and tend to
 CC have a unique polypyrrolone type II extended helix structure that permits
 CC them to span the membrane with relatively few amino acids. The peptides
 CC possess the ability to work synergistically with antibiotics, and in
 CC addition, some of them possess anti-endotoxin activity.
 CC N.B. The present sequence represents SEQ ID NO:1 in the claims and
 CC examples of the specification, but differs slightly from the SEQ ID NO:1
 CC in the sequence listing on page 51 of the specification (see AAW27179).

XX
 SQ Sequence 13 AA;

Query Match 38.6%; Score 91; DB 18; Length 13;
 Best Local Similarity 100.0%; Pred. NO. 4e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24 ILKKPMPMPWRK 36
 Db 1 ILKKPMPMPWRK 13

RESULT 15

AAW24609
 ID AAW24609 standard; peptide; 13 AA.

XX
 AC AAW24609;

DT 18-AUG-1999 (first entry)

DE Indolicidin analogue #61.

KW Indolicidin; bacterial infection; photo-oxidised solubiliser;

KW antimicrobial; antibiotic; antiarrhythmic; surface disinfectant;

KW additive; shampoo; soap; insecticide; herbicide; preservative;

XX food; technical material.

OS Synthetic.

XX WO9807745-A2.

PD 26-FEB-1998.

PF 21-AUG-1997; 97WO-US14779.

PR 13-JAN-1997; 97US-0034949.

XX 21-AUG-1996; 96US-0024754.

PA (MIGR-) MICROLOGIX BIOTECH INC.

PI Ernie D, Fraser JR, Krieger TJ, Taylor R, West MH;

XX WPI; 1998-169090/15.

PT New indolicidin analogues with antimicrobial activity and related
 PT nucleic acid - vectors, transformed cells and antibodies, also
 PT conjugates with polyoxalkylene glycol and fatty acid to reduce

PT toxicity, useful therapeutically, as disinfectants etc.

XX
 PS Example 1; Page 32; 129pp; English.

XX
 CC AAW24549 to AAW24615 represent indolicidin analogues of formulae
 CC (I)-(VIII) containing up to 25 amino acids (aa): $RX_2XX_2XB(1)$, BX_2XX_2XB
 CC (II), $BBX_2XX_2XB(111)$, $BX_2XX_2XB(Bn(AA)Nm)LBAGS(IV)$, $BX_2XX_2XB(AA)Nm$
 CC (V), $LBnX_2XX_2XB(111)$, $LBnX_2XX_2XB(111)$ and $BX_2XX_2XB(111)$.
 CC Where $Z = P$ or V ; $X =$ hydrophobic residue, preferably W; B = basic aa;
 CC preferably R or K; AA = any aa; $n = 0$ or 1; in (II), at least 1 Z = V;
 CC in (VIII) at least 2 X = F or Y. The analogues are used to treat
 CC infections caused by bacteria (Gram positive or negative, or anaerobic);
 CC fungi (yeast or moulds); parasites (protozoa, nematodes, cestodes or
 CC trematodes) or viruses. Typical of very many pathogens that can be
 CC controlled are *Leishmania*, *Trypanosoma*, *Ascaris lumbricoides*, *Fasciola*
 CC *hepatica*, *Klebsiella pneumoniae*, *Bordetella pertussis*, *Staphylococcus*
 CC *aureus*, *Listeria*, *Clostridium*, *rotavirus* and *papilloma virus*. Compounds
 CC derived from the analogues may be used similarly; the compounds may
 CC also be prepared from antibiotics or antiarrhythmic agents. The analogues
 CC may be used therapeutically or to coat medical devices; also they are
 CC useful as surface disinfectants, as additives to shampoo or soaps, as
 CC insecticides or herbicides, or as preservatives for foods and technical
 CC materials. The analogues are administered by injection, lavage, orally
 CC or topically, generally at 0.1-50 mg/kg. These analogues have a broader
 CC spectrum of activity than indolicidin and modification as compounds
 CC reduces their toxicity.

XX
 SQ Sequence 13 AA;

Query Match 38.6%; Score 91; DB 19; Length 13;
 Best Local Similarity 100.0%; Pred. NO. 4e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24 ILKKPMPMPWRK 36
 Db 1 ILKKPMPMPWRK 13

Search completed: July 29, 2003, 16:31:41
 Job time : 59.6849 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2003, 16:29:40 ; Search time 20.274 Seconds
(without alignments)
77.217 Million cell updates/sec

Title: SEQ35-27-35
Perfect score: 241
Sequence: 1 ILKKPMPWRRKHEAPEEPIMLKKPMPWRRK 37

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PTCUS.COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/Dackfilest.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query length	DB ID	Description
1	141.5	58.7	28	4	US-09-030-619-104 Sequence 104, App
2	137.5	57.1	28	4	US-09-030-619-50 Sequence 50, Appl
3	133.5	55.4	63	4	US-09-099-631A-12 Sequence 12, Appl
4	133.5	55.4	63	4	US-09-416-481A-39 Sequence 39, Appl
5	99.5	41.3	16	3	US-08-702-054B-38 Sequence 38, Appl
6	99.5	41.1	21	3	US-08-915-314-54 Sequence 54, Appl
7	99.5	41.1	21	4	US-09-030-619-69 Sequence 69, Appl
8	99.5	41.1	21	4	US-09-667-486-54 Sequence 54, Appl
9	98.5	40.7	21	3	US-08-915-314-56 Sequence 56, Appl
10	98.5	40.7	21	3	US-09-030-619-52 Sequence 52, Appl
11	98.5	40.7	21	4	US-09-667-486-55 Sequence 55, Appl
12	93.5	38.8	20	3	US-08-915-314-55 Sequence 55, Appl
13	93.5	38.8	20	4	US-09-030-619-51 Sequence 51, Appl
14	93.5	38.8	20	4	US-09-667-486-55 Sequence 55, Appl
15	91.5	37.8	13	3	US-08-915-314-30 Sequence 30, Appl
16	91.5	37.8	13	3	US-08-915-314-62 Sequence 62, Appl
17	91.5	37.8	13	3	US-08-915-314-63 Sequence 63, Appl
18	91.5	37.8	13	3	US-08-915-314-64 Sequence 64, Appl
19	91.5	37.8	13	3	US-09-042-071-36 Sequence 36, Appl
20	91.5	37.8	13	4	US-09-030-619-95 Sequence 95, Appl
21	91.5	37.8	13	4	US-09-030-619-99 Sequence 99, Appl
22	91.5	37.8	13	4	US-09-667-486-30 Sequence 30, Appl
23	91.5	37.8	13	4	US-09-667-486-62 Sequence 62, Appl
24	91.5	37.8	13	4	US-09-667-486-63 Sequence 63, Appl
25	91.5	37.8	13	4	US-09-667-486-64 Sequence 64, Appl
26	91.5	37.8	14	4	US-08-915-314-57 Sequence 57, Appl
27	91.5	37.8	14	4	US-09-030-619-72 Sequence 72, Appl

28	91	37.8	14	4	US-09-030-619-108 Sequence 108, App
29	91	37.8	15	4	US-09-667-486-57 Sequence 57, Appl
30	89	36.9	14	3	US-08-702-054B-40 Sequence 40, Appl
31	87.5	36.3	21	3	US-08-915-314-46 Sequence 46, Appl
32	87.5	36.3	21	4	US-09-030-619-47 Sequence 47, Appl
33	87.5	36.3	21	4	US-09-667-486-46 Sequence 46, Appl
34	87	36.1	12	3	US-08-915-314-52 Sequence 52, Appl
35	87	36.1	12	4	US-09-030-619-67 Sequence 67, Appl
36	87	36.1	12	4	US-09-667-486-52 Sequence 52, Appl
37	86	35.7	12	3	US-08-915-314-74 Sequence 74, Appl
38	86	35.7	12	3	US-08-702-054B-5 Sequence 5, Appl
39	86	35.7	12	4	US-09-030-619-112 Sequence 112, App
40	86	35.7	12	4	US-09-667-486-74 Sequence 74, Appl
41	86	35.7	13	3	US-08-915-314-51 Sequence 51, Appl
42	86	35.7	13	3	US-08-915-314-58 Sequence 58, Appl
43	86	35.7	13	3	US-08-702-054B-34 Sequence 34, Appl
44	86	35.7	13	4	US-09-030-619-53 Sequence 53, Appl
45	86	35.7	13	4	US-09-030-619-107 Sequence 107, App

ALIGNMENTS

RESULT 1
US-09-030-619-104
Sequence 104, Application US/09030619B
Patent No. 6503881
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfile, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 104
LENGTH: 28
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-104
Query Match 58.7%; Score 141.5; DB 4; Length 28;
Best Local Similarity 58.8%; Pred. No. 1.6e-11;
Matches 20; Conservative 2; Mismatches 1; Indels 11; Gaps 1;
QY 4 KMPMPWRRKHEAPEEPIMLKKPMPWRRK 37
DB 3 KMPMPWRRK-----ILMKPMPWRRK 25
RESULT 2
US-09-030-619-50
Sequence 50, Application US/09030619B
Patent No. 6503881
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfile, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
WITH ANTIBIOTICS

```
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030.619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO: 50
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-50

Query Match
Best Local Similarity 57.1%; Score 137.5; DB 4; Length 28;
Matches 21; Conservative 2; Mismatches 0; Indels 11; Gaps 2;

OY 4 KWPMPWRRKHAEPEAEPIIMLK-KWPMPWRRK 37
    :|||||||  |||:|||||||
Db 3 RWPMPWRRK-----MIL-KWPMPWRRK 25

RESULT 3
US-09-099-631A-12
; Sequence 12, Application US/09099631A
; Patent No. 6444645
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; TITLE OF INVENTION: Crosslink-Stabilized Indolicidin Analogs
; FILE REFERENCE: P-UC 3050
; CURRENT APPLICATION NUMBER: US/09/099,631A
; CURRENT FILING DATE: 1998-06-18
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 12
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-099-631A-12

Query Match
Best Local Similarity 55.4%; Score 133.5; DB 4; Length 63;
Matches 22; Conservative 0; Mismatches 5; Indels 7; Gaps 2;

OY 4 KWPMPWRRKHAEPEAEPIIMLK-KWPMPWRR 36
    :|||||||  |||:|||||||
Db 11 KWPMPWRRM-----ARIAMILPKWMPWRR 38

RESULT 4
US-09-416-481A-39
; Sequence 39, Application US/09416481A
; Patent No. 6524585
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; TITLE OF INVENTION: Indolicidin Analogs and Methods of Using Same
; FILE REFERENCE: P-UC 3794
; CURRENT APPLICATION NUMBER: US/09/416,481A
; CURRENT FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: US 09/076,227
; PRIOR FILING DATE: 1998-05-12
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 39
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
```

```
US-09-416-481A-39

Query Match
Best Local Similarity 55.4%; Score 133.5; DB 4; Length 63;
Matches 22; Conservative 0; Mismatches 5; Indels 7; Gaps 2;

OY 4 KWPMPWRRKHAEPEAEPIIMLK-KWPMPWRR 36
    :|||||||  |||:|||||||
Db 11 KWPMPWRRM-----ARIAMILPKWMPWRR 38

RESULT 5
US-08-702-054B-38
; Sequence 38, Application US/08702054B
; Patent No. 6191254
; GENERAL INFORMATION:
; APPLICANT: Falls, Timothy J.
; APPLICANT: Hancock, Robert E. W.
; APPLICANT: Gough, Monisha
; TITLE OF INVENTION: ANTIMICROBIAL CATIONIC PEPTIDES
; TITLE OF INVENTION: AND METHODS OF SCREENING FOR THE SAME
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,054B
; FILING DATE: 23-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/002,687
; FILING DATE: 23-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Halle, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07420/013001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-702-054B-38

Query Match
Best Local Similarity 41.3%; Score 99.5; DB 3; Length 16;
Matches 16; Conservative 0; Mismatches 0; Indels 21; Gaps 1;

OY 1 ILKKWMPWRRKHAEPEAEPIIMLKWMPWRRK 37
    :|||||||
Db 1 ILKKWMPW-----WMPWMPWRRK 16

RESULT 6
US-08-915-314-54
; Sequence 54, Application US/08915314
; Patent No. 6180604
; GENERAL INFORMATION:
; APPLICANT: Frazer, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
```

APPLICANT: Erfle, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear

US-08-915-314-54

Query Match 41.1%; Score 99; DB 3; Length 21;
Best Local Similarity 64.3%; Pred. No. 2.4e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 ILKKPMPWPKRKHAEPEAPIMILKK 28
DB 1 ILKKPMPWPKRKHAEPEAPIMILKK 18

RESULT 7
US-09-030-619-69

Sequence 69, Application US/09030619B
Patent No. 6503861

GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfle, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
TITLE OF INVENTION: WITH ANTIBIOTICS
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 69

LENGTH: 21
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Indolicidin Analogue

US-09-030-619-69

Query Match 41.1%; Score 99; DB 4; Length 21;
Best Local Similarity 64.3%; Pred. No. 2.4e-06;

Matches 18; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 ILKKPMPWPKRKHAEPEAPIMILKK 28
DB 1 ILKKPMPWPKRKHAEPEAPIMILKK 18

RESULT 8
US-09-667-486-54

Sequence 54, Application US/09667486
Patent No. 6538106

GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfle, Douglas

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/667,486
FILING DATE: 22-Sep-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997

ATTORNEY/AGENT INFORMATION:
NAME: No. 6538106tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 54:
US-09-667-486-54

Query Match 41.1%; Score 99; DB 4; Length 21;
Best Local Similarity 64.3%; Pred. No. 2.4e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 ILKKPMPWPKRKHAEPEAPIMILKK 28
DB 1 ILKKPMPWPKRKHAEPEAPIMILKK 18

RESULT 9
US-08-915-314-56

Sequence 56, Application US/08915314
Patent No. 6180604

GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert

APPLICANT: Erfile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tendburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-56

Query Match 40.7%; Score 98; DB 3; Length 21;
Best Local Similarity 64.3%; Pred. No. 3.2e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 ILKKPMPWRRRHEAPEAPIMILKK 28
Db 1 ILKKPMPWRRR-----IMILKK 18

RESULT 10
US-09-030-619-52
Sequence 52, Application US/09030619B
Patent No. 6503881
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfile, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,619B
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 52
LENGTH: 21
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-52

Query Match 40.7%; Score 98; DB 4; Length 21;
Best Local Similarity 64.3%; Pred. No. 3.2e-06;

Matches 18; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 ILKKPMPWRRRHEAPEAPIMILKK 28
Db 1 ILKKPMPWRRR-----IMILKK 18

RESULT 11
US-09-667-486-56
Sequence 56, Application US/09667486
Patent No. 6538106
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/667,486
FILING DATE: 22-SEP-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: No. 6538106tendburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 56:
US-09-667-486-56

Query Match 40.7%; Score 98; DB 4; Length 21;
Best Local Similarity 64.3%; Pred. No. 3.2e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 ILKKPMPWRRRHEAPEAPIMILKK 28
Db 1 ILKKPMPWRRR-----IMILKK 18

RESULT 12
US-08-915-314-55
Sequence 55, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert

APPLICANT: Erfle, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-55

Query Match 38.8%; Score 93.5; DB 3: Length 20;
Best Local Similarity 60.7%; Pred. No. 1.1e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 11; Gaps 1;

Db 1 ILKKPMPWPMRRKHEAPEAPIMILKK 28
1 ILKKPMPWPMRR-----MILKK 17

RESULT 13
US-09-030-619-51
Sequence 51, Application US/09030619B
Patent No. 6503881
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfle, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 51
LENGTH: 20
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-51

Query Match 38.8%; Score 93.5; DB 4: Length 20;
Best Local Similarity 60.7%; Pred. No. 1.1e-05;

Matches 17; Conservative 0; Mismatches 0; Indels 11; Gaps 1;

Db 1 ILKKPMPWPMRRKHEAPEAPIMILKK 28
1 ILKKPMPWPMRR-----MILKK 17

RESULT 14
US-09-667-486-55
Sequence 55, Application US/09667486
Patent No. 6538106
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfle, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/667,486
FILING DATE: 22-SEP-2000
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: No. 6538106tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: <unknown>
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 55:
US-09-667-486-55

Query Match 38.8%; Score 93.5; DB 4: Length 20;
Best Local Similarity 60.7%; Pred. No. 1.1e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 11; Gaps 1;

Db 1 ILKKPMPWPMRRKHEAPEAPIMILKK 28
1 ILKKPMPWPMRR-----MILKK 17

RESULT 15
US-08-915-314-30
Sequence 30, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert

```

APPLICANT: Erfile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-30

```

```

Query Match      37.8%; Score 91; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0;

```

```

OY      1 ILKKWPMWPMRRK 13
        |||
Db      1 ILKKWPMWPMRRK 13

```

```

Search completed: July 29, 2003, 16:36:06
Job time : 20.274 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2003, 16:28:25 ; Search time 24.8356 Seconds

(without alignments)
143.272 Million cell updates/sec

Title: SEQ35-27-35

Perfect score: 241

Sequence: 1 ILKKMPMPMPRRKHEAPEAPIMILKKMPMPMPRRK 37

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	32.4	144	1	JC1222
2	63	26.1	192	2	H86543
3	63	26.1	192	2	D72081
4	62	25.7	314	2	S43916
5	60	24.9	381	2	B87470
6	59.5	24.7	114	2	T36208
7	59.5	24.7	2290	1	GNMYE
8	59	24.5	95	2	E86447
9	59	24.5	485	2	S74708
10	58	24.1	107	2	T35634
11	58	24.1	513	2	S21976
12	57.5	23.9	376	2	T40591
13	57	23.7	711	2	C40046
14	56	23.2	452	2	T28094
15	55.5	23.0	527	2	S33068
16	55.5	23.0	990	1	G46335
17	55.5	23.0	1940	1	A59287
18	55	22.8	245	2	C69459
19	55	22.8	352	2	S77448
20	55	22.8	715	2	B70741
21	54.5	22.6	134	2	E72532
22	54.5	22.6	340	2	PH0217
23	54.5	22.6	451	2	S30401
24	54.5	22.6	547	2	T45635
25	54	22.4	144	2	S35331
26	54	22.4	187	2	AB1647
27	54	22.4	469	2	B70607
28	54	22.4	473	2	C86949
29	54	22.4	1173	1	VG1HHC

30	53.5	22.2	538	2	B84759	hypothetical prote
31	53.5	22.2	1113	2	JE0315	low-density lipopr
32	53	22.0	68	2	E82799	hypothetical prote
33	53	22.0	406	2	H69143	coenzyme F420-redu
34	53	22.0	424	2	T07742	omega-6 desaturase
35	53	22.0	480	2	JC7552	Shb-1-like adapter p
36	53	22.0	691	2	D71430	hypothetical prote
37	53	22.0	949	2	E75352	glycine cleavage s
38	53	22.0	2292	2	S35961	capsid polypeptid
39	52.5	21.8	35	2	T11026	H+-transporting tw
40	52.5	21.8	228	2	AD3326	chloramphenicol O-
41	52.5	21.8	449	2	C84618	hypothetical prote
42	52.5	21.8	471	2	T50016	transcriptional fact
43	52.5	21.8	490	2	T21365	hypothetical prote
44	52.5	21.8	518	2	T29589	hypothetical prote
45	52.5	21.8	621	2	S37664	peptidomeric polyp

ALIGNMENTS

```

RESULT 1
JC1222
Indolicidin precursor - bovine
N:Alternate names: antimicrobial peptide
C:Species: Bos primigenius taurus (cattle)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: JC1222; A42387; S25664
R:del Sal, G.; Storici, P.; Schneider, C.; Romeo, D.; Zanetti, M.
Biochem. Biophys. Res. Commun. 187, 467-472, 1992
A:Title: cDNA cloning of the neutrophil bactericidal peptide indolicidin.
A:Reference number: JC1222; MUID:92392368; PMID:1520337
A:Accession: JC1222
A:Molecule type: mRNA
A:Residues: 1-144 <SALD>
A:Cross-references: EMBL:X67340; NID:g462; PIDN:CAA47755.1; PID:g463
A:Experimental source: bone marrow
R:Selsted, M.E.; Novotny, M.J.; Morris, W.L.; Tang, Y.Q.; Smith, W.; Cullor, J.S.
J. Biol. Chem. 267, 4292-4295, 1992
A:Title: Indolicidin, a novel bactericidal tripeptide amide from neutrophils.
A:Reference number: A42387; MUID:92165771; PMID:1537821
A:Accession: A42387
A:Molecule type: protein
A:Residues: 131-143 <SELD>
A:Experimental source: neutrophils
A:Note: sequence extracted from NCBI backbone (NCBIP:83840)
C:Superfamily: cathelin; cystatin homology
C:Keywords: amidated carboxyl end
F:1-29/Domain: signal sequence #status predicted <SIG>
F:22-129/Domain: cystatin homology <CYS>
F:30-130/Domain: propeptide #status predicted <PRO>
F:131-143/Product: indolicidin #status experimental <MAT>
F:143/Modified site: amidated carboxyl end (Arg) (amide in mature form from following
Query Match 32.4%; Score 78; DB 1; Length 144;
Best local similarity 55.6%; Pred. No. 0.011;
Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 19 EAEPIMILKKMPMPMPRR 36
DB 126 EHQSVILLPWKMPMPMPRR 143
RESULT 2
H86543
hypothetical protein CPJ0426 [imported] - Chlamydia pneumoniae (strain J138)
C:Species: Chlamydia pneumoniae; Chlamydia pneumoniae
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C:Accession: H86543
C:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362

```

A:Accession: H86543
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-192 <STO>
 A:Cross-references: GB:BA000008; NID:98978798; PIDN:BA98634.1; GSPDB:GN00142
 C:Genetics:
 A:Gene: CPJ0426

Query Match 26.1%; Score 63; DB 2; Length 192;
 Best Local Similarity 42.3%; Pred. No. 1;
 Matches 11; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 5 WPMWPMRRKHAEPEAPIMILKKMP 30
 Db 138 WPMPLPKRKQIEKLPKGEICPLSAYP 163

RESULT 3

D72081
 conserved hypothetical protein frameshifted CP0327 [imported] - Chlamydia pneumoniae
 N:Alternate names: hypothetical protein CR277 homolog
 C:Species: Chlamydia pneumoniae
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
 C:Accession: D72081; G81589

R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
 Nature Genet. 21, 385-389, 1999

A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
 A:Reference number: A72000; MUID:99206606; PMID:10192388

A:Accession: D72081
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-192 <ARN>

A:Cross-references: GB:AE001625; GB:AE001363; NID:94376695; PIDN:AA018570.1; PID:9437670
 A:Experimental source: strain CMT029

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
 C.; Dodson, R.; Gwin, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg,
 Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.
 A:Reference number: A81500; MUID:20150255; PMID:10684935

A:Accession: G81589
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-192 <RNA>

A:Cross-references: GB:AE002195; GB:AE002161; NID:97189246; PIDN:AA38182.1; PID:9718925
 A:Experimental source: strain AR39, HL cells

C:Genetics:
 A:Gene: CP0426; CP0327

Query Match 26.1%; Score 63; DB 2; Length 192;
 Best Local Similarity 42.3%; Pred. No. 1;
 Matches 11; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 5 WPMWPMRRKHAEPEAPIMILKKMP 30
 Db 138 WPMPLPKRKQIEKLPKGEICPLSAYP 163

RESULT 4

S43916
 hypothetical protein 3 - Bacillus stearothermophilus
 C:Species: Bacillus stearothermophilus
 C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 15-Oct-1999
 C:Accession: S43916

R:Long, E.C.; Ferenci, T.;
 Mol. Gen. Genet. 243, 343-352, 1994

A:Title: Molecular cloning of a maltose transport gene from Bacillus stearothermophilus
 A:Reference number: S43914; MUID:94247374; PMID:8130087

A:Accession: S43916
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-314 <LTO>

A:Cross-references: EMBL:L13418; NID:9436964; PIDN:AA71981.1; PID:9436967
 A:Note: the authors translated the initiation codon GCG for residue 1 as Val

Query Match 25.7%; Score 62; DB 2; Length 314;
 Best Local Similarity 34.7%; Pred. No. 2.3;
 Matches 17; Conservative 5; Mismatches 13; Indels 14; Gaps 3;

QY 2 LKKWPMWPMRRKHAEPEA-----EP-----IMILKKWPMW-----WRR 36
 Db 222 LKOWMEQAKARWEAEARDADGARRPGETWDFLLDKPSWWRLLGRRMR 270

RESULT 5

B87470
 hypothetical protein CC1782 [imported] - Caulobacter crescentus
 C:Species: Caulobacter crescentus
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
 C:Accession: B87470

R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
 B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Halt, D.H.; Ko
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.
 A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: B87470
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-381 <STO>

A:Cross-references: GB:AE005673; NID:913423210; PIDN:AAK23758.1; GSPDB:GN00148
 C:Genetics:
 A:Gene: CC1782

Query Match 24.9%; Score 60; DB 2; Length 381;
 Best Local Similarity 37.0%; Pred. No. 5;
 Matches 10; Conservative 4; Mismatches 9; Indels 4; Gaps 1;

QY 7 WMP-----WRRKHAEPEAPIMILKKW 29
 Db 170 WMPRRGMRKALDKMKRAAPLRKLYDM 196

RESULT 6

T36208
 hypothetical protein SCE36.09 - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T36208

R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A
 Submitted to the EMBL Data Library, May 1999

A:Reference number: 221601
 A:Accession: T36208

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA

A:Residues: 1114 <COLI>

A:Cross-references: EMBL:AL049763; PIDN:CAB42078.1; GSPDB:GN00070; SCOEDB:SCE36.09
 A:Experimental source: strain A3(2)

C:Genetics:
 A:Gene: SCOEDB:SCE36.09

Query Match 24.7%; Score 59.5; DB 2; Length 114;
 Best Local Similarity 50.0%; Pred. No. 1.6;
 Matches 10; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

QY 18 PEAPIMILKKW-PWMPMR 36
 Db 93 PETAPADARAKWRPMPMR 112

RESULT 7

GNNYE
 genome polyprotein - encephalomyocarditis virus
 N:Contains: coat protein VP1; coat protein VP2; coat protein VP3; coat protein VP4; c
 EC 3.4.4.1; RNA-directed RNA polymerase (EC 2.7.7.48)

C:Species: encephalomyocarditis virus, EMCV
 A:Note: host Homo sapiens (man)

C>Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 16-Jul-1999
 C:Accession: A03906; JN0383
 R:Palmerberg, A.C.; Kirby, E.M.; Janda, M.R.; Drake, N.L.; Duke, G.M.; Potratz, K.F.; CC
 Nucleic Acids Res. 12, 2969-2985, 1984
 A>Title: The nucleotide and deduced amino acid sequences of the encephalomyocarditis vir
 A:Reference number: A03906; MUID:84169586; PMID:6324136
 A:Accession: A03906
 A:Molecule type: genomic RNA
 A:Residues: 1-2290 <PAL>
 A:Cross-references: GB:X00463; MID:g61034; PIDN:CA25152.1; PID:g61035
 R:Petkov, N.A.; Chizhikov, V.E.; Blinov, V.M.; Karginov, V.A.; Mikryukov, N.N.; Gutorov,
 Bioorg. Khim. 10, 274-279, 1984
 A>Title: Nucleotide sequence of the 3'-terminus of encephalomyocarditis virus RNA.
 A:Reference number: JN0383; MUID:85022788; PMID:6091680
 A:Accession: JN0383
 A:Molecule type: genomic RNA
 A:Residues: 1357-1396; 'L', 1398-1517, 'A', 1519-1536, 'E', 1538-1556, 'S', 1558-1611, 'T', 1613-1
 A:Cross-references: GB:M54935
 A>Note: the authors translated the codon CAU for residue 713 as Thr and AAC for residue
 C:Superfamily: foot-and-mouth disease virus genome polypeptide
 C:Keywords: coat protein; core protein; genome-linked protein; hydrolase; nucleotidyltra
 F:1-67/Domain: leader peptide #status predicted <LDP>
 F:68-136/Product: coat protein VP4 #status predicted <VP4>
 F:137-391/Product: coat protein VP2 #status predicted <VP2>
 F:392-622/Product: coat protein VP3 #status predicted <VP3>
 F:623-910/Product: coat protein VP1 #status predicted <VP1>
 F:911-1056/Product: core protein P2-A #status predicted <P2A>
 F:1057-1192/Product: core protein P2-B #status predicted <P2B>
 F:1193-1517/Product: core protein P2-C #status predicted <P2C>
 F:1518-1605/Product: core protein P3-A #status predicted <P3A>
 F:1606-1625/Product: genome-linked protein VPg #status predicted <VPg>
 F:1626-1830/Product: proteinase #status predicted <PTS>
 F:1831-2290/Product: RNA-directed RNA polymerase #status predicted <RNP>

Query Match 24.7%; Score 59.5; DB 1; Length 2290;
 Best Local Similarity 26.2%; Pred. No. 38;
 Matches 11; Conservative 6; Mismatches 8; Indels 17; Gaps 1;

QY 6 PMWPMRRKHEAPEPEP-----MILKKP 30
 Db 967 PMWPMKNTYQAVLRAEPCRVMTDLYKKRPPRLPLVQKEWP 1008

RESULT 8
 E86447
 Protein F5D14.5 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: E86447
 R:Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Hultar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C:J.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: E86447
 A:Molecule type: preliminary
 A:Status: preliminary
 A:Residues: 1-95 <STO>
 A:Cross-references: GB:AE005172; MID:g8920603; PIDN:AAE81325.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F5D14.5
 A:Map position: 1

Query Match 24.5%; Score 59; DB 2; Length 95;
 Best Local Similarity 21.3%; Pred. No. 1.6;
 Matches 10; Conservative 5; Mismatches 4; Indels 28; Gaps 3;

QY 5 WP-----WPMRRKHEAPEPEMILK----KPMWPM 34
 Db 46 WPMVYVAVGAGGGRMMW-----PVLITDVGGGSMWMM 81

RESULT 9
 S74708
 hypothetical protein slr1306 - Synechocystis sp. (strain PCC 6803)
 C:Species: Synechocystis sp.
 A:Variety: PCC 6803
 C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
 C:Accession: S74708
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,
 O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
 DNA Res. 3, 109-116, 1996
 A>Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys
 S.
 A:Reference number: S74322; MUID:97061201; PMID:8905231
 A:Accession: S74708
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-485 <KAN>
 A:Cross-references: EMBL:D90901; GB:AB001339; MID:g1651897; PIDN:BA16859.1; PID:d101
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 24.5%; Score 59; DB 2; Length 485;
 Best Local Similarity 29.3%; Pred. No. 8.6;
 Matches 12; Conservative 6; Mismatches 5; Indels 18; Gaps 2;

QY 6 PM-----WPMRRKHEAPEPEM---ILKK 28
 Db 41 PMDWGLMALSGLVIRWRRRRHAPPEQOMLTPREVLOK 81

RESULT 10
 T35634
 hypothetical protein SC6G9.38 - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
 C:Accession: T35634
 R:Seeger, K.J.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A
 submitted to the EMBL Data Library, June 1999
 A:Reference number: 221584
 A:Accession: T35634
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-107 <SEE>
 A:Cross-references: EMBL:AL079356; PIDN:CAB45629.1; GSPDB:GN00070; SCOPDB:SC6G9.38
 C:Genetics:
 A:Experimental source: strain A3(2)
 A:Gene: SCOPDB:SC6G9.38

Query Match 24.1%; Score 58; DB 2; Length 107;
 Best Local Similarity 50.0%; Pred. No. 2.3;
 Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 3 KPMWPMRRKHEAPEPEP 22
 Db 79 RKWPMKPERPSAAPEADTP 98

RESULT 11
 S21976
 probable RNA-directed DNA polymerase (EC 2.7.7.49) (clone MH2C) - rat retrotransposon
 N:Alternate names: reverse transcriptase
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 19-Mar-1997 #sequence_revision 17-Oct-1997 #text_change 26-Aug-1999
 C:Accession: S21976
 R:Kahre, O.; Ilyes, H.; Speck, M.
 submitted to the EMBL Data Library, August 1991
 A:Reference number: S16783
 A:Accession: S21976
 A:Molecule type: mRNA

A:Residues: 1-513 <KAH>
 A:Cross-references: EMBL:X61295; NID:g56521; PIDN:CAA43593.1; PID:g56522
 A:Experimental source: clone MH2C
 C:Genetics:
 A:Mobile element: retrotransposon L1
 C:Superfamily: pol polyprotein
 C:Keywords: nucleotidyltransferase; polyprotein; reverse transcriptase

Query Match 24.1%; Score 58; DB 2; Length 513;
 Best Local Similarity 39.4%; Pred. No. 12;
 Matches 13; Conservative 4; Mismatches 13; Indels 4; Gaps 2;

OY 1 ILKKPWPWPWR--RKHEPEAEPIMLIK-KW 29
 | | | | | : : : : : | | | | |
 DB 373 IFSKWCWFNRATCRHQIDPOLSPCTKLKSKW 405

RESULT 12

T40591

hypothetical protein SPBC646.15c - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 04-Mar-2000

C:Accession: T40591

R:Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, January 1999

A:Reference number: 221938

A:Accession: T40591

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-376 <SEE>

A:Cross-references: EMBL:AL035216; PIDN:CAA22819.1; GSPDB:GN00067; SPDB:SPBC646.15c

A:Experimental source: strain 972h-; cosmid c646

C:Genetics:

A:Gene: SPDB:SPBC646.15c

A:Map position: 2

A:Introns: 49/71; 216/2; 312/2; 350/1

C:Superfamily: Schizosaccharomyces pombe hypothetical protein SPBC646.15c

Query Match 23.9%; Score 57.5; DB 2; Length 376;
 Best Local Similarity 37.0%; Pred. No. 10;
 Matches 10; Conservative 3; Mismatches 7; Indels 7; Gaps 1;

OY 8 WPRRRKHEAPEEPIMLIKWPWPW 34
 | | | | | : : : : : | | | | |
 DB 236 WHMKRKOKSSS-----LKVFPWGPW 255

RESULT 13

C40046

antibiotic transport-associated protein actII-3 - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 22-Oct-1999

C:Accession: C40046

R:Fernandez-Moreno, M.A.; Caballero, J.L.; Hopwood, D.A.; Malpartida, F.

A:Title: The act cluster contains regulatory and antibiotic export genes, direct targets

A:Reference number: A40046; MUID:91347376; PMID:1878971

A:Accession: C40046

A:Molecule type: DNA

A:Residues: 1-711 <FER>

A:Cross-references: GB:M64683; NID:g153143; PIDN:AAA26691.1; PID:g153146

Query Match 23.7%; Score 57; DB 2; Length 711;
 Best Local Similarity 47.4%; Pred. No. 23;
 Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 1 ILKKPWPWPWRKHEAPEE 19
 | | | | | : : : : : | | | | |
 DB 332 IFGRWVFWPARPKHTEPD 350

RESULT 14

T28094

hypothetical protein ZK899.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000

C:Accession: T28094

R:Kershaw, J.

submitted to the EMBL Data Library, September 1994

A:Reference number: Z20468

A:Accession: T28094

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-452 <WIL>

A:Cross-references: EMBL:Z37140; PIDN:CAA85502.1; GSPDB:GN00028; CESP:ZK899.2

A:Experimental source: clone ZK899

C:Genetics:

A:Gene: CESP:ZK899.2

A:Map position: X

A:Introns: 34/3; 143/2; 227/2; 262/3; 380/3

C:Superfamily: Caenorhabditis elegans hypothetical protein ZK899.2

Query Match 23.2%; Score 56; DB 2; Length 452;
 Best Local Similarity 32.3%; Pred. No. 19;
 Matches 10; Conservative 3; Mismatches 10; Indels 8; Gaps 2;

OY 4 KWPWPWPWRKHEAPEEPIMLIKWPWPW 34
 | | | | | : : : : : | | | | |
 DB 172 KLVWWTW---HDTDPN-----IFDRMNWVPW 194

RESULT 15

S33068

myosin heavy chain - fluke (Schistosoma mansoni) (fragment)

N:Alternate names: surface antigen, 200K

C:Species: Schistosoma mansoni

C>Date: 22-Nov-1993 #sequence_revision 06-Sep-1996 #text_change 13-Feb-1998

C:Accession: S33068

R:Solomon, L.M.A.; Masterson, C.P.; Tom, T.D.; McNally, M.T.; Lowell, G.H.; Strand, M

J.; Immunol. 149, 3612-3620, 1992

A:Title: Induction of protective immunity in mice using a 62-kDa recombinant fragment

A:Reference number: A46514; MUID:93056536; PMID:1431131

A:Accession: S33068

A:Molecule type: mRNA

A:Residues: 1-527 <SOI>

A:Cross-references: EMBL:X65591

A>Note: the authors translated the codon CAA for residue 346 as Lys

C:Superfamily: myosin heavy chain; myosin motor domain homology

C:Keywords: ATP; surface antigen

Query Match 23.0%; Score 55.5; DB 2; Length 527;
 Best Local Similarity 30.8%; Pred. No. 25;
 Matches 12; Conservative 7; Mismatches 9; Indels 11; Gaps 2;

OY 1 ILKKPWPW-----PW----RKHEAPEEPIMLIK 28
 | | | | | : : : : : | | | | |
 DB 106 VLRNWPWPRLYTKVPMLNIAROEEEMKKAEBELAKKE 144

Search completed: July 29, 2003, 16:35:20
 Job time : 26.8356 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2003, 16:19:05 ; Search time 13.6849 Seconds

(without alignments)
127.146 Million cell updates/sec

Title: SEQ35-27-35
Perfect score: 241
Sequence: 1 ILKKPMPWMPRRKHEAPEPIMILKKMPWMPRRK 37

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_A1:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	32.4	144	1	INDC_BOVIN
2	62	25.7	314	1	YMA3_BACST
3	59.5	24.7	2290	1	POLG_EMCV
4	57	23.7	711	1	MMLA_STRCO
5	55.5	23.0	55	1	ATP8_ANAPL
6	55.5	23.0	942	1	ENV_CAENG
7	55.5	23.0	990	1	ENV_OMVVS
8	55	22.8	715	1	YD55_MYCTU
9	54	22.4	144	1	AP22_APIME
10	54	22.4	469	1	SYCL_MYCLE
11	54	22.4	473	1	SYCL_MYCLE
12	54	22.4	1173	1	VG12_CVH22
13	53.5	22.2	1042	1	CORI_HUMAN
14	53.5	22.2	1113	1	CORI_MOUSE
15	53	22.0	424	1	FP6C_SOYBN
16	52.5	21.8	55	1	ATP8_AYTMN
17	52.5	21.8	257	1	E434_ADE40
18	52.5	21.8	691	1	YH1G_ECOLI
19	52.5	21.8	982	1	ENV_VILV
20	52.5	21.8	983	1	ENV_VILV
21	52.5	21.8	991	1	ENV_VILV
22	52.5	21.8	1154	1	VG12_IBVD2
23	52.5	21.8	1162	1	VG12_IBVD2
24	52.5	21.8	1162	1	VG12_IBVD2
25	52.5	21.8	1162	1	VG12_IBVD2
26	52.5	21.8	1163	1	VG12_IBVD2
27	52	21.6	68	1	Y121_BP74
28	51.5	21.4	162	1	DSBB_NEIMA
29	51.5	21.4	162	1	DSBB_NEIMA
30	51.5	21.4	175	1	NGGC_ANASP
31	51.5	21.4	295	1	CHO_ECOLI
32	51.5	21.4	443	1	FP6C_BRANA
33	51.5	21.4	448	1	FP6C_ARATH

ALIGNMENTS

RESULT 1	ID	INDC_BOVIN	STANDARD:	PRT:	144 AA.
AC	P33046:				
DT	01-OCT-1993 (Rel. 27, Created)				
DT	01-OCT-1993 (Rel. 27, Last sequence update)				
DE	15-SEP-2003 (Rel. 42, Last annotation update)				
OS	Indolicidin precursor.				
OS	Bos taurus (Bovine).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
OC	Bovidae; Bovinae; Bos.				
OX	NCBI_TaxID=9913;				
RM	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE-Bone marrow;				
RC	MEDLINE=92392368; PubMed=1520337;				
RA	del Sal G., Storici P., Schneider C., Romeo D., Zanetti M.;				
RT	"cDNA cloning of the neutrophil bactericidal peptide indolicidin."				
RL	Biochem. Biophys. Res. Commun. 187:467-472(1992).				
RM	[2]				
RP	SEQUENCE OF 131-143.				
RC	TISSUE-Neutrophils;				
RC	MEDLINE=92165771; PubMed=1537821;				
RA	Selsted M.E., Novotny M.J., Morris W.L., Tang Y.-Q., Smith W.;				
RT	Cullor J.S.;				
RT	"Indolicidin, a novel bactericidal tridecapeptide amide from				
RL	neutrophils."				
CC	J. Biol. Chem. 267:4292-4295(1992).				
CC	- FUNCTION: POTENT MICROBICIDAL ACTIVITY, ACTIVE AGAINST				
CC	STAPHYLOCOCCUS AUREUS AND ESCHERICHIA COLI.				
CC	- TISSUE SPECIFICITY: LARGE GRANULES OF NEUTROPHILS.				
CC	- PTM: ELASTASE MIGHT BE RESPONSIBLE FOR ITS MATURATION.				
CC	- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
CC	EMBL: X67340; CAA47755.1; -				
CC	PIR: JCI1222; JCI1222.				
CC	PDB: 1G89; 17-JAN-01.				
CC	PDB: 1G8C; 17-JAN-01.				
CC	PDB: 1H81; 31-DEC-02.				
CC	InterPro: IPR001894; Cathelicidin.				
CC	Pfam: PF00666; Cathelicidins; 1.				
CC	ProDom: PD001838; Cathelicidins; 1.				
CC	PROSITE: PS00946; CATHELICIDINS_1; 1.				
CC	PROSITE: PS00947; CATHELICIDINS_2; 1.				
CC	Antibiotic: Amidation; Signal; Pyroglutamate carboxylic acid;				
CC	3D-structure.				
FT	SIGNAL 1 29 POTENTIAL.				

34 51.5 21.4 989 1 ENV_VILV1 P23422 visna lentii
35 51.5 21.4 1262 1 MYO6_HUMAN O9um54 homo sapien
36 51.5 21.4 1265 1 MYO6_MOUSE O64331 mus musculu
37 51 21.2 126 1 YD43_MYCTU O11013 mycobacteri
38 51 21.2 196 1 YAO5_SCHPO O09677 schizosach
39 51 21.2 295 1 CHO_ECO57 O8xd27 escherichia
40 51 21.2 295 1 CHO_ECOLI P76213 escherichia
41 51 21.2 295 1 CHO_SHIFL P59361 shigella fl
42 51 21.2 556 1 MEND_ECOLI P17109 e menaquin
43 51 21.2 2292 1 POLG_EMCVB P17593 encephalomy
44 51 21.2 2292 1 POLG_EMCVB P17594 encephalomy
45 50.5 21.0 293 1 CHO_SALTI O82695 salmonella

```

FT PROPEP 30 130
FT PEPTIDE 131 143
FT MOD_RES 30 30
FT DISULFID 85 96
FT DISULFID 107 124
FT MOD_RES 143 143
SQ SEQUENCE 144 AA: 16479 MW: 531CBHE55C09911 CRC64;

Query Match
Best Local Similarity 32.4%; Score 78; DB 1; Length 144;
Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 19 EAEPIILIKKWPMPWR 36
DB 126 ELQSVILPKWMPWR 143

RESULT 2
YMA3_BACST
ID YMA3_BACST STANDARD; PRT; 314 AA.
AC Q45633:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 35.7 kDa protein in mala 3' region (ORF3).
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
OX NCBI_Taxid=1422;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 7953;
RX MEDLINE=94247374; PubMed=8190087;
RA Liong E.C., Ferenc T.;
RT "Molecular cloning of a maltose transport gene from Bacillus
RT stearothermophilus and its expression in Escherichia coli K-12.";
RL Mol. Gen. Genet. 243:343-352(1994).
CC -1- SIMILARITY: BELONGS TO THE UPF0097 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L13418; AAA71981.1;
DR PIR: S43916;
DR InterPro: IPR004175; 2_5_ligase.
DR Pfam: PF02834; 2_5_ligase; 2.
KW Hypothetical protein.
SQ SEQUENCE 314 AA: 35735 MW: B54E25FD3F72BFC4 CRC64;

Query Match
Best Local Similarity 25.7%; Score 62; DB 1; Length 314;
Matches 17; Conservative 5; Mismatches 13; Indels 14; Gaps 3;

QY 2 LKKWPMWPKRRKHEAPEA-----EP-----IMILKKWPMW-----WRR 36
DB 222 LKQWMDQAKARWEADADGAREPGETWDFLLDKPSWWRILGRWR 270

RESULT 3
POLG_EMCV
ID POLG_EMCV STANDARD; PRT; 2290 AA.
AC P03304:
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 to VP4; Core proteins
DE P2A to P2C; P3A; Genome-linked protein VPG; Picornain 3C
DE (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D
```

```

DE (EC 2.7.7.48)).
OS Encephalomyocarditis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Cardiovirus.
OX NCBI_Taxid=12104;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84169586; PubMed=6324136;
RA Palmenberg A.C., Kirby E.M., Janda M.R., Drake N.L., Duke G.M.,
RA Potratz K.F., Collett M.S.;
RT "The nucleotide and deduced amino acid sequences of the
RT encephalomyocarditis viral polyprotein coding region.";
RL Nucleic Acids Res. 12:2969-2985(1984).
CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC O/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC -1- CATALYTIC ACTIVITY: Selective cleavage of Gln-1-Gly bond in the
CC polyovirus polyprotein. In other picornavirus reactions Gln may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X00463; CA25152.1;
DR PIR: A03906; GNNVE.
DR HSSP: P12296; 2MEV.
DR MEROPS: C03.009;
DR InterPro: IPR001676; Rhv.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSV1r.
DR Pfam: PF00073; Rhv; 3.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
KW polyprotein; Coat protein; Core protein; Transferrase;
KW RNA-directed RNA polymerase; Hydrolyase; Thiol protease; Myristate.
FT PROPEP 1
FT CHAIN 68 136
FT CHAIN 137 391
FT CHAIN 392 622
FT CHAIN 623 910
FT CHAIN 911 1056
FT CHAIN 1057 1192
FT CHAIN 1193 1517
FT CHAIN 1518 1605
FT CHAIN 1606 1625
FT CHAIN 1626 1830
FT CHAIN 1831 2290
FT LIPID 68 68
FT ACT_SITE 1784 1784
FT ACT_SITE 1802 1802
SQ SEQUENCE 2290 AA: 255756 MW: 20BC81B7CF08C85 CRC64;

Query Match
Best Local Similarity 24.7%; Score 59.5; DB 1; Length 2290;
Matches 11; Conservative 6; Mismatches 8; Indels 17; Gaps 1;

QY 6 PWWPMWRKRRKHEAPEA-----IMILKKW 30
DB 967 PWWPMWRKRRKHEAPEA-----IMILKKW 1008
```


SEQUENCE	4	MLA_STRCO	STANDARD	PRT	711 AA
ID	MLA_STRCO				
AC	Q35902				
DT	30-MAY-2000 (Rel. 39, Created)				
DT	30-MAY-2000 (Rel. 39, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Putative membrane protein actii-3				
OS	ACTII-3 OR SCOS084 OR SCBAC28G1.10.				
OC	Streptomyces coelicolor.				
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;				
OC	Streptomycinae; Streptomycetaceae; Streptomycetes.				
OX	NCBI_Taxid=1902;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	MEDLINE=91347376; PubMed=1878971;				
RA	Fernandez-Moreno M.A., Caballero J.L., Hopwood D.A., Malpartida F.;				
RT	"The act cluster contains regulatory and antibiotic export genes,				
RT	direct targets for translational control by the bldA tRNA gene of				
RT	Streptomyces.";				
RL	Cell 66:769-780(1991).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=AS(2) / M145;				
RA	MEDLINE=AS(2) / M145;				
RA	Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,				
RA	Thompson N.R., James K.D., Harris D.E., Quail M.A., Kiser H.,				
RA	Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,				
RA	Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,				
RA	Huang C.-H., Kiser J., Laake L., Murphy L., Oliver K., O'Neill S.,				
RA	Rabblnowitsch E., Rajandream M.A., Rutherford K., Rutter S.,				
RA	Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,				
RA	Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,				
RA	Hopwood D.A.;				
RT	"Complete genome sequence of the model actinomycete Streptomyces				
RT	coelicolor AS(2).";				
RL	Nature 417:141-147(2002).				
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).				
CC	-1- SIMILARITY: BELONGS TO THE MPL FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation				
CC	the European Bioinformatics Institute. There are no restrictions on				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL: M64683; AAA26691.1; -				
DR	EMBL: AL939122; CAC44197.1; -				
DR	PIR: C40046; C40046.				
DR	InterPro: IPR004707; ActII.				
DR	InterPro: IPR004869; MPL.				
DR	InterPro: IPR000731; SSD_5TM.				
DR	Pfam: PF03176; MPL; 2.				
DR	TIGRFAMS: TIGR00833; actII; 1.				
DR	PROSITE: PSS0156; SSD; 2.				
KW	Hypothetical protein; Transmembrane; Complete proteome.				
FT	TRANSMEM 14				
FT	TRANSMEM 175				
FT	TRANSMEM 199				
FT	TRANSMEM 235				
FT	TRANSMEM 281				
FT	TRANSMEM 313				
FT	TRANSMEM 369				
FT	TRANSMEM 516				
FT	TRANSMEM 540				
FT	TRANSMEM 573				
FT	TRANSMEM 623				
FT	TRANSMEM 645				
FT	SEQUENCE 711 AA; 74862 MW; AS460BDEABED1B6 CRC64;				

01	Query Match	23.7%	Score 57;	DB 1;	Length 711;
02	Best Local Similarity	47.4%	Pred. No. 11;		
03	Matches	9;	Conservative	3;	Mismatches 7; Indels 0; Gaps 0;
04					
05	Db	332	IFGHWVFWPAPRKHGTEDP	350	
06					
07					
08					
09					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26					
27					
28					
29					
30					
31					
32					
33					
34					
35					
36					
37					
38					
39					
40					
41					
42					
43					
44					
45					
46					
47					
48					
49					
50					
51					
52					
53					
54					
55					
56					
57					
58					
59					
60					
61					
62					
63					
64					
65					
66					
67					
68					
69					
70					
71					
72					
73					
74					
75					
76					
77					
78					
79					
80					
81					
82					
83					
84					
85					
86					
87					
88					
89					
90					
91					
92					
93					
94					
95					
96					
97					
98					
99					
100					

```
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11662;
RN [1]
RP SEQUENCE FROM N.A. AND SEQUENCE OF 81-95.
RX MEDLINE=92015464; PubMed=1656067;
RA Knowles D.P. Jr., Cheevers W.P., McGuire T.C., Brassfield A.L.,
RA Harwood W.G., Stem T.A.;
RT "Structure and genetic variability of envelope glycoproteins of two
RT antigenic variants of caprine arthritis-encephalitis lentivirus.";
RU J. Virol. 63:5744-5750(1991).
RN [2]
RP REVISIONS.
RA Knowles D.P.;
RU Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

EMBL: M60855; AAB88709.2; -
InterPro: IPR000328; Env_Gp41.
DR Pfam: PF00517; Gp41; 1.
KW Glycoprotein; Coat protein; Polyprotein; Transmembrane.
FT PEPTIDE 1 80
FT CHAIN 81 630
FT CHAIN 631 942
FT DOMAIN 1 630
FT TRANSMEMBRANE PROTEIN (POTENTIAL).
FT DOMAIN 631 659
FT TRANSMEM 660 799
FT TRANSMEM 800 820
FT DOMAIN 821 942
FT CHAIN 51 51
FT CARBOHYD 98 98
FT CARBOHYD 131 131
FT CARBOHYD 176 176
FT CARBOHYD 228 228
FT CARBOHYD 331 331
FT CARBOHYD 348 348
FT CARBOHYD 354 354
FT CARBOHYD 370 370
FT CARBOHYD 379 379
FT CARBOHYD 400 400
FT CARBOHYD 404 404
FT CARBOHYD 435 435
FT CARBOHYD 441 441
FT CARBOHYD 447 447
FT CARBOHYD 457 457
FT CARBOHYD 467 467
FT CARBOHYD 481 481
FT CARBOHYD 493 493
FT CARBOHYD 503 503
FT CARBOHYD 509 509
FT CARBOHYD 527 527
FT CARBOHYD 534 534
SEQUENCE 942 AA: 108437 MW: 5808150355F4A CRC64;

Query Match 23.0%; Score 55.5; DB 1; Length 942;
Best Local Similarity 21.4%; Pred. No. 22;
Matches 15; Conservative 7; Mismatches 9; Indels 39; Gaps 4;

QY 3 KKPWP-----WP-----WRKHAEPEAE-----PIMI 25
DB 143 ENWPNTYHWPLOQENYRWIKENIAENKKRKNSTKKGIELLAGTIRGRVCPYPAL 202
QY 26 LK--KWPWP 33
DB 203 LKTKWCWP 212

```
RESULT 7
ENV_OMVVS
ID ENV_OMVVS STANDARD: PRT: 990 AA.
AC P16899:
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE ENV polyprotein precursor (Coat polyprotein).
GN ENV.
OS Ovine lentivirus (strain SA-OMV).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11664;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90223989; PubMed=2158181;
RA Querat G., Audoly G., Sonigo P., Vigne R.;
RT "Nucleotide sequence analysis of SA-OMV, a virus-related ovine
RT lentivirus: phylogenetic history of lentiviruses.";
RU Virology 175:434-447(1990).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

EMBL: M34193; AAA6783.1; -
EMBL: M31646; AAA6817.1; -
DR HIV: M34193; ENVSONVVSACG.
DR InterPro: IPR000328; Env_Gp41.
DR Pfam: PF00517; Gp41; 1.
KW Glycoprotein; Coat protein; Polyprotein; Transmembrane.
FT PEPTIDE 1 101
FT CHAIN 102 662
FT CHAIN 663 990
FT TRANSMEM 842 863
FT CARBOHYD 141 141
FT CARBOHYD 162 162
FT CARBOHYD 207 207
FT CARBOHYD 259 259
FT CARBOHYD 299 299
FT CARBOHYD 363 363
FT CARBOHYD 386 386
FT CARBOHYD 402 402
FT CARBOHYD 413 413
FT CARBOHYD 434 434
FT CARBOHYD 438 438
FT CARBOHYD 469 469
FT CARBOHYD 474 474
FT CARBOHYD 480 480
FT CARBOHYD 490 490
FT CARBOHYD 500 500
FT CARBOHYD 514 514
FT CARBOHYD 526 526
FT CARBOHYD 536 536
FT CARBOHYD 542 542
FT CARBOHYD 550 550
FT CARBOHYD 560 560
FT CARBOHYD 567 567
FT CARBOHYD 703 703
FT CARBOHYD 771 771
FT CARBOHYD 778 778
FT CARBOHYD 794 794
SEQUENCE 990 AA: 114498 MW: 279881685561F3 CRC64;

Query Match 23.0%; Score 55.5; DB 1; Length 990;
Best Local Similarity 20.0%; Pred. No. 23;
Matches 14; Conservative 9; Mismatches 8; Indels 39; Gaps 4;

QY 3 KKPWP-----WP-----WRKHAEPEAE-----PIMI 25

```

Db      174 QEWPNNTYHWPIDMKNRNMKNENEKEYTSNNKTKEDIDALLAGKINGRCVYPPAL 233
Oy      26 LK--KMPMP 33
        || :| :|
Db      234 LKCEMCWYP 243

RESULT 8
ID      YD55_MYCTU      STANDARD:      PRT:      715 AA.
AC      011025:
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Hypothetical protein RV1355C.
GN      RV1355C OR MT1398 OR MTCY02B10.19C.
OS      Mycobacterium tuberculosis.
OC      Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC      Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX      NCBI_TaxID=1773;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-H37RV:
RX      MEDLINE=96295987; PubMed=9634230;
RA      Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA      Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA      Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA      Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA      Hornsby T., Jorgels K., Krogh A., McLean J., Moule S., Murphy L.,
RA      Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA      Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA      Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT      "Deciphering the biology of Mycobacterium tuberculosis from the
RT      complete genome sequence."
RL      Nature 393:537-544(1998).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=CDC 1551 / Oshkosh;
RA      Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA      Peterson J., Deboy R., Dodson R., Gwin M.L., Haft D., Hickey E.,
RA      Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA      Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA      Bishai W.;
RT      "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT      laboratory strains."
RL      Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; 275555; CAA99988.1;
DR      EMBL; AE007012; AAK45661.1; ALT_INIT.
DR      PIR; B70741; B70741.
DR      TIGR; MT1398;
DR      Tuberclust; RV1355C;
DR      InterPro; IPR000594; Th1F_domain.
DR      Pfam; PF00899; Th1F; 1.
KW      Hypothetical protein; Complete proteome.
SQ      SEQUENCE 715 AA; 78181 MW; 455495248A56041C CRC64;

Query Match      22.8%; Score 55; DB 1; Length 715;
Best Local Similarity 38.1%; Pred. No. 19;
Matches      8; Conservative 7; Mismatches 4; Indels 2; Gaps 1;

Oy      18 PEAEPIIM-LKKMPMPMR 36
        | :| :| :| :| :|
Db      53 POPDDLEBAKRWAYIFMR 73

```

```

RESULT 9
ID      AP22_APIME      STANDARD:      PRT:      144 AA.
AC      P3581; P11525; P11526;
DT      01-OCT-1989 (Rel. 12, Created)
DT      01-JUN-1994 (Rel. 29, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Apidaecin precursor, type 22.
OS      Apis mellifera (honeybee).
OC      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC      Neoptera; Apis.
OC      Apidae; Apis.
OX      NCBI_TaxID=7460;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      MEDLINE=93223697; PubMed=8467807;
RA      Casteels-Josson K., Capaci T., Casteels P., Tempst P.;
RT      "Apidaecin multipetide precursor structure: a putative mechanism for
RT      amplification of the insect antibacterial response."
RL      EMBO J. 12:1569-1578(1993).
RN      [2]
RP      SEQUENCE (APIDAEICIN IA/IB).
RC      TISSUE=Hemolymph;
RX      MEDLINE=90005446; PubMed=2676519;
RA      Casteels P., Ampe C., Jacobs F., Vaeck M., Tempst P.;
RT      "Apidaecins: antibacterial peptides from honeybees."
RL      EMBO J. 8:2387-2391(1989).
CC      -1- FUNCTION: APIDAEICIN HAVE BACTERICIDAL ACTIVITY; PREDOMINANTLY
CC      AGAINST GRAM-NEGATIVE BACTERIA. THEY SEEM TO INTERFERE WITH CELL
CC      PROLIFERATION.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X72576; CAA51168.1;
DR      PIR; S35331; S35331.
DR      InterPro; IPR004828; Apidaecin.
DR      Pfam; PF00807; Apidaecin; 4.
KW      Insect immunity; Antibiotic; Hemolymph; Signal; Multigene family;
KW      Cleavage on pair of basic residues; Repeat.
FT      SIGNAL 1 19
FT      PROPEP 20 42
FT      PEPTIDE 43 60
FT      PROPEP 63 70
FT      PEPTIDE 71 88
FT      PROPEP 91 98
FT      PEPTIDE 99 116
FT      PROPEP 119 126
FT      PEPTIDE 127 144
SQ      SEQUENCE 144 AA; 16539 MW; 6FA1AD74C677108D CRC64;

Query Match      22.4%; Score 54; DB 1; Length 144;
Best Local Similarity 64.7%; Pred. No. 5.4;
Matches      11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Oy      6 PMPWRKHEAPEAP 22
        | :| :| :| :| :|
Db      26 PRRPRLRREAREAP 42

RESULT 10
ID      SYCL_MYCTU      STANDARD:      PRT:      469 AA.
AC      P96862;
DT      15-JUL-1998 (Rel. 36, Created)
DT      15-JUL-1998 (Rel. 36, Last sequence update)
DT      15-SEP-2003 (Rel. 42, Last annotation update)

```

Dd	89	PWMEWMATHERAFTAAVDALDVLPSSAEP	117
RESULT	11		
ID	SYCL_MYCLE	STANDARD:	PRT; 473 AA.
AC	p5790;		
DT	16-OCT-2001 (Rel. 40, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	15-SEP-2003 (Rel. 42, Last annotation update)		
DE	CysteinyI-tRNA synthetase 1 (EC 6.1.1.16) (Cysteine--tRNA ligase 1) (CysRS 1).		
DE	(CysRS 1).		
OS	CYSYL OR CYSS OR ML0323.		
OS	Mycobacterium leprae.		
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;		
OC	Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.		
OX	NCBI_TaxID=1769;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	STRAIN-TN.		
RA	Medline-21128732; PubMed-11234002;		
RA	Cole S.T., Eigleier K., Parkhill J., James K.D., Thomson N.R., Wheeler P.R., Honore N., Garnier H., Churcher C., Harris D., Mungall K., Basham D., Brown D., Chillingworth T., Connor R., Davies R.M., Deakin K., Duthoy S., Felwell T., Fraser A., Hamlin N., Holtroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S., Murphy L., Oliver K., Quail M.A., Rajadream M.A., Rutherford K.M., Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R., Barrell B.G.,		
RT	"Massive gene decay in the leprosy bacillus."		
RL	Nature 409:1007-1011(2001).		
CC	-1- CATALYTIC ACTIVITY: ATP + L-cysteine + tRNA(Cys) = AMP +		
CC	-1- diphosphate + L-cysteinyl-tRNA(Cys).		
CC	-1- SUBUNIT: Monomer (By similarity).		
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic.		
CC	-1- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family. Strong, to methionyl-tRNA synthetase.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).		
CC	-----		
CC	EMBL: AL583918; GAC29831.1; -		
CC	PIR: C86949; C86949.		
DR	Leptoma; ML0323; -		
DR	HAWAP; MF_00041; -; 1.		
DR	InterPro; IPR002308; Cys_tRNA-synt_1a.		
DR	InterPro; IPR001412; tRNA-synt_1.		
DR	Pfam; PF01406; tRNA-synt_1e; 1.		
DR	PRINTS; PR00983; TRNASYNTHCS.		
DR	TIGRFAMS; TIGR00435; CysS; 1.		
DR	PROSITE; PS00178; AA_TRNA_LIGASE_I; FALSE_NEG.		
KM	Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding; Complete proteome.		
KW	Complete proteome.		
FT	Site 35 45 "HIGH" REGION.		
FT	Site 267 271 "KMSKS" REGION.		
FT	Binding 270 270 ATP (By Similarity).		
SO	Sequence 473 AA; 52124 MW; 9FDGCF859C69316D CRC64;		
Oy	Query Match 22.4%; Score 54; DB 1; Length 473; Best Local Similarity 34.5%; Pred. No. 17; Matches 10; Conservative 0; Mismatches 7; Indels 12; Gaps 1;		
Oy	6 PWMWRKKHE-----AEPEKEP 22 		
Oy	89 PWMEWMATHERAFTAAVDALDVLPSSAEP 117		

RESULT 12
VGL2_CVH22 STANDARD: PRT: 1173 AA
ID VGL2_CVH22
AC P15423; P89342; P89343; P89344; Q6174; Q990M1; Q990M2; Q990M3;
Q990M4;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE E2 glycoprotein precursor (Spike glycoprotein) (Peplomer protein).
GN S.
OS Human coronavirus (strain 229E) (HCoV-229E).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=11137;
RN RN
RP [1]
RX MEDLINE FROM N.A.
RX MEDLINE=90264837; PubMed=2345367;
RA Raabe T., Schelle-Prinz B., Siddell S.G.;
RT "Nucleotide sequence of the gene encoding the spike glycoprotein of
RL human coronavirus HCoV 229E".
RT J. Gen. Virol. 71:1065-1073(1990).
RN [2]
RP MEDLINE FROM N.A.
RX MEDLINE=21262210; PubMed=11369870;
RA Thiel V., Herold J., Schelle B., Siddell S.G.;
RT "Infectious RNA transcribed in vitro from a cDNA copy of the human
RL coronavirus genome cloned in vaccinia virus.";
RT J. Gen. Virol. 82:1273-1281(2001).
RN [3]
RP MEDLINE FROM N.A.
RX MEDLINE=21262210; PubMed=11369870;
RA Thiel V., Herold J., Schelle B., Siddell S.G.;
RT "Infectious RNA transcribed in vitro from a cDNA copy of the human
RL coronavirus genome cloned in vaccinia virus.";
RT J. Gen. Virol. 82:1273-1281(2001).
RN [4]
RP MEDLINE=9086140; PubMed=9870593;
RX STRAIN=Isolate ATCC VR-74, Isolate A162, and Isolate LRI 281;
RA STRAIN=99086140; PubMed=9870593;
RT Hays J.P., Myint S.H.;
RT "Viral and cellular changes in a human cell line persistently infected
RL with human coronavirus HCoV-229E.";
RN Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
RN [5]
RP MEDLINE=9086140; PubMed=9870593;
RX STRAIN=Isolate ATCC VR-74, Isolate A162, and Isolate LRI 281;
RA STRAIN=99086140; PubMed=9870593;
RT Hays J.P., Myint S.H.;
RT "PCR sequencing of the spike genes of geographically and
RL chronologically distinct human coronaviruses 229E.";
RT J. Virol. Methods 75:179-193(1998).
RN [6]
RP MEDLINE=9086140; PubMed=9870593;
RX STRAIN=Isolate ATCC VR-74, Isolate A162, and Isolate LRI 281;
RA STRAIN=99086140; PubMed=9870593;
RT Hays J.P., Myint S.H.;
RT "PCR sequencing of the spike genes of geographically and
RL chronologically distinct human coronaviruses 229E.";
RT J. Virol. Methods 75:179-193(1998).
RN [7]
RP MEDLINE=22440020; PubMed=1255191;
RX Bonavia A., Zelus B.D., Wentworth D.E., Talbot P.J., Holmes K.V.;
RT "Identification of a receptor-binding domain of the spike glycoprotein
RL of human coronavirus HCoV-229E.";
RT J. Virol. 77:2530-2538(2003).
RN [8]
RP MEDLINE=22521439; PubMed=12634402;
RX Breslin J.J., Mork I., Smith M.K., Vogel L.K., Hemmila E.M.,
RA Bonavia A., Talbot P.J., Steciw H., Noren O., Holmes K.V.;
RT "Human coronavirus 229E: Receptor binding domain and neutralization by
RL soluble receptor at 37 degrees C.";
RT J. Virol. 77:4435-4438(2003).
RN [9]
RP MEDLINE=21109095; PubMed=11162792;
RX Gallagher T.M., Buchmeier M.J.;
RT "Coronavirus spike proteins in viral entry and pathogenesis.";
RL Virology 279:371-374(2001).

CC	-1	FUNCTION: Structural protein that makes spikes at the surface of the virus. Determines enteropathogenicity and virulence of the virus. Initiates infection by specifically recognizing and binding the human aminopeptidase ANPEP receptor. Its association with ANPEP may lead to its conformational change that triggers fusion between viral and host cellular membrane.
CC	-1	SUBUNIT: Homotrimer. During virus morphogenesis, it is found in a complex with M and HE proteins (By similarity). Interacts with ANPEP.
CC	-1	SUBCELLULAR LOCATION: Type I membrane protein.
CC	-1	DOMAIN: The spike S1 domain displays the specificity for the host receptor.
CC	-1	DOMAIN: The leucine zipper-like heptad repeats may mediate the fusion of viral and cellular membranes.
CC	-1	POLYMORPHISM: The strong variation between the different strains may affect the virulence of the virus.
CC	-1	MISCELLANEOUS: In contrast to serogroup 2, E2 glycoprotein protein from serogroup 1 is not cleaved.
CC	-1	SIMILARITY: Contains 1 spike S1 domain.
CC	-1	SIMILARITY: Contains 1 spike S2 domain.
CC	-1	-----
CC	-1	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).
CC	-1	-----
DR	EMBL:	X16816; CAA34723.1; -
DR	EMBL:	AF304460; AAG48592.1; -
DR	EMBL:	AF344186; AAK32188.1; -
DR	EMBL:	AF344187; AAK32189.1; -
DR	EMBL:	AF344188; AAK32190.1; -
DR	EMBL:	AF344189; AAK32191.1; -
DR	EMBL:	Y09923; CAA71056.1; -
DR	EMBL:	Y10051; CAA71146.1; -
DR	EMBL:	Y10052; CAA71147.1; -
DR	EMBL:	X15654; CAA33680.1; -
DR	PIR:	A34766; VGIHHC.
DR	InterPro:	IPRO02551; Corona_S1.
DR	InterPro:	IPRO02552; Corona_S2.
DR	Pfam:	PF01600; Corona_S1; 1.
KW	Pfam:	PF01601; Corona_S2; 1.
KM	Virulence:	Glycoprotein; Envelope protein; Transmembrane; Signal;
KV	Collided coll.	
FT	SIGNAL	1 15
FT	CHAIN	16 1173
FT	DOMAIN	16 1115 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	1116 1135 POTENTIAL.
FT	DOMAIN	1136 1173 CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	32 536 SPIKE S1.
FT	DOMAIN	417 547 INTERACTION WITH ANPEP.
FT	DOMAIN	537 1171 SPIKE S2.
FT	DOMAIN	1054 1103 COLLED COIL (POTENTIAL).
FT	DOMAIN	1067 1102 LEUCINE ZIPPER-LIKE HEPATAD REPEATS.
FT	DOMAIN	1136 1157 CYS-RICH.
FT	CARBOHYD	23 23 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	171 171 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	220 220 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	243 243 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	326 326 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	440 440 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	464 464 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	518 518 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	538 538 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	542 542 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	568 568 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

FT CARBOHYD 581 581 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 587 587 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 663 663 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 930 930 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1015 1015 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1020 1020 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1037 1037 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1049 1049 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1061 1061 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1066 1066 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1076 1076 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1082 1082 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1096 1096 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 98 98 N -> S (in isolate LRI 281).
FT VARIANT 120 120 N -> I (in isolate LRI 281).
FT VARIANT 127 127 LR -> IS (in isolate A162).
FT VARIANT 176 176 N -> T (in isolate P100E).
FT VARIANT 210 210 T -> S (in isolate A162).
FT VARIANT 223 223 T -> N (in isolate A162).
FT VARIANT 228 228 DF -> V (in isolate A162).
FT VARIANT 230 230 C -> L (in isolate LRI 281).
FT VARIANT 230 230 C -> F (in isolates RW Stock, P11A, P11B,
FT VARIANT 230 230 P100E and ATCC VR-74).
FT VARIANT 248 248 S -> A (in isolate A162).
FT VARIANT 270 270 D -> V (in isolate P100E).
FT VARIANT 295 295 V -> A (in isolate LRI 281).
FT VARIANT 300 300 T -> M (in isolate P100E).
FT VARIANT 307 307 D -> N (in isolate A162).
FT VARIANT 311 311 PQ -> LR (in isolate A162).
FT VARIANT 314 314 GGCRCNCPAG -> VGRICNCPAV (in isolate
FT VARIANT 336 336 K -> N (in isolate LRI 281).
FT VARIANT 349 336 KYVANYANGV -> QVYAKFD (in isolate A162).
FT VARIANT 401 401 V -> M (in isolate A162).
FT VARIANT 404 404 WAYSRYT -> LANNSHN (in isolate A162).
FT VARIANT 414 414 S -> T (in isolate P100E).
FT VARIANT 424 424 G -> V (in isolate A162).
FT VARIANT 430 430 O -> K (in isolate A162).
FT VARIANT 441 441 V -> A (in isolate LRI 281).
FT VARIANT 444 444 D -> N (in isolate A162).
FT VARIANT 462 462 V -> I (in isolate A162).
FT VARIANT 481 481 L -> V (in isolate A162).
FT VARIANT 488 488 K -> N (in isolate A162).
FT VARIANT 530 530 L -> M (in isolate A162).
FT VARIANT 577 577 I -> T (in isolate P11A).
FT VARIANT 578 578 V -> G (in isolate P11B).
FT VARIANT 590 590 T -> I (in isolate P100E).
FT VARIANT 642 642 T -> R (in isolate A162).
FT VARIANT 681 681 R -> R (in isolate A162).
FT VARIANT 700 700 L -> I (in isolates RW Stock, P11A, P11B
FT VARIANT 711 711 D -> N (in isolate LRI 281).
FT VARIANT 714 714 K -> N (in isolates RW Stock, P11A, P11B
FT VARIANT 765 765 V -> A (in isolate A162).
FT VARIANT 775 775 A -> S (in isolate A162).

Query Match 22.4%; Score 54; DB 1; Length 1173;
Best Local Similarity 85.7%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Oy 4 KWPMPW 10
Db 1113 KWPMPW 1119

RESULT 13
CORI_HUMAN STANDARD; PRT; 1042 AA.
AC Q9Y5Q5:Q9YH2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

```

```

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Atrial natriuretic peptide-converting enzyme (EC 3.4.21.-) (Pro-ANP-
DE Converting enzyme) (Corin) (Heart specific serine proteinase ATC2).
GN CRN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=99262646; PubMed=10329693;
RA Yan W., Sheng N., Soto M., Morser J., Wu Q.;
RT "Corin, a mosaic transmembrane serine protease encoded by a novel cDNA
RT from human heart."
RL Eur. J. Biol. Chem. 274:14926-14935(1999).
RN [2]
RP SEQUENCE OF 734-1040 FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=20354769; PubMed=11082206;
RA Hooper J.D., Scarnan A.L., Clarke B.E., Normyle J.F., Antalis T.M.;
RT "Localization of the mosaic transmembrane serine protease corin to
RT heart myocytes."
RL Eur. J. Biochem. 267:6931-6937(2000).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=20359740; PubMed=1080574;
RA Yan W., Wu F., Morser J., Wu Q.;
RT "Corin, a transmembrane cardiac serine protease, acts as a pro-atrial
RT natriuretic peptide-converting enzyme."
RL Proc. Natl. Acad. Sci. U.S.A. 97:8525-8529(2000).
CC -I- FUNCTION: CONVERTS PRO-ANP TO ANP. CLEAVES PRO-ANP SPECIFICALLY
CC BETWEEN ARG-123 AND SER-124.
CC -I- SUBCELLULAR LOCATION: Type II membrane protein.
CC -I- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN HEART. EXPRESSED IN HEART
CC MYOCYTES.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -I- SIMILARITY: Contains 7 LDL-receptor class A domains.
CC -I- SIMILARITY: Contains 2 tritzled (F2) domains.
CC -I- SIMILARITY: Contains 1 SRCR domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AF133845; AAD31850.1; -;
DR EMBL, AF113246; AAF21966.1; -;
DR HSP; P00763; IDPO.
DR MEROPS; S01.019; -.
DR MIM; 605236; -.
DR GO; GO:0005887; C: integral to plasma membrane; TAS.
DR GO; GO:0007345; P: embryogenesis and morphogenesis; TAS.
DR GO; GO:0006629; P: lipid metabolism; TAS.
DR GO; GO:0006508; P: proteolysis and peptidolysis; TAS.
DR GO; GO:0008217; P: regulation of blood pressure; TAS.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000024; Fz domain.
DR InterPro: IPR002172; LDL_receptor_A.
DR InterPro: IPR001254; Ser_protease_TY.
DR InterPro: IPR001190; Srcr_receptor.
DR Pfam; PF01392; Fz; 2.
DR Pfam; PF00057; Idl_recept_a; 6.
DR Pfam; PF00089; trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00063; FRI; 2.
DR SMART; SM00192; LDLa; 3.
DR SMART; SM00202; SR; 1.
DR SMART; SM00020; Tryp_Spc; 1.

```

```

DR PROSITE: PS50038; FZ; 2.
DR PROSITE: PS01209; LDLRA_1; 6.
DR PROSITE: PS50068; LDLRA_2; 7.
DR PROSITE: PS50240; TRYPsin_DOM; 1.
DR PROSITE: PS00134; TRYPsin_HIS; FALSE_NEG.
DR PROSITE: PS00135; TRYPsin_SER; 1.
DR PROSITE: PS00420; SRCR_1; FALSE_NEG.
DR PROSITE: PS50287; SRCR_2; FALSE_NEG.
KW Hydrolyase; Serine protease; Transmembrane; Signal-anchor;
  Glycoprotein; Repeat.
FT DOMAIN 1 45
FT TRANSMEM 46 66
FT
FT DOMAIN 67 1042
FT DOMAIN 134 259
FT DOMAIN 268 304
FT DOMAIN 305 340
FT DOMAIN 341 377
FT DOMAIN 378 415
FT DOMAIN 450 573
FT DOMAIN 579 614
FT DOMAIN 615 653
FT DOMAIN 654 690
FT DOMAIN 690 786
FT DOMAIN 802 1042
FT ACT_SITE 843 843
FT ACT_SITE 892 892
FT ACT_SITE 985 985
FT ACT_SITE 985 985
FT DISULFID 790 912
FT DISULFID 828 844
FT DISULFID 955 970
FT DISULFID 981 1010
FT CARBOHYD 80 80
FT CARBOHYD 104 104
FT CARBOHYD 135 135
FT CARBOHYD 141 141
FT CARBOHYD 231 231
FT CARBOHYD 245 245
FT CARBOHYD 251 251
FT CARBOHYD 305 305
FT CARBOHYD 320 320
FT CARBOHYD 376 376
FT CARBOHYD 413 413
FT CARBOHYD 446 446
FT CARBOHYD 451 451
FT CARBOHYD 469 469
FT CARBOHYD 567 567
FT CARBOHYD 651 651
FT CARBOHYD 697 697
FT CARBOHYD 761 761
FT CARBOHYD 1022 1022
FT MUTAGEN 985 985
FT CONFLICT 854 854
FT CONFLICT 876 876
SO SEQUENCE 1042 AA; 116564 MW; 7705398EBB07AD2 CRC64;

Query Match 22.2%; Score 53.5; DB 1; length 1042;
Best Local Similarity 36.0%; Pred. No. 42;
Matches 9; Conservative 5; Mismatches 8; Indels 3; Gaps 1;

```

```

DE converting enzyme) (Corin) (Low density lipoprotein receptor related
DE protein 4).
GN CRN OR LRP4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-98429596; PubMed-9756624;
RA Tomita Y., Kim D.-H., Magoori K., Fujino T., Yamamoto T.T.;
RT "A novel low-density lipoprotein receptor-related protein with type II
RT membrane protein-like structure is abundant in heart.";
RL J. Biochem. 124:784-789(1998).
CC
CC -I- FUNCTION: CONVERTS PRO-ANP TO ANP. CLEAVES PRO-ANP SPECIFICALLY
CC BETWEEN ARG-122 AND SER-123 (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: Type II membrane protein.
CC -I- TISSUE SPECIFICITY: Highly expressed in heart.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -I- SIMILARITY: Contains 7 LDL-receptor class A domains.
CC -I- SIMILARITY: Contains 2 fritzled (FZ) domains.
CC -I- SIMILARITY: Contains 1 SRCR domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AB013874; BAA34371.1; -
CC PIR: J00315; J00315.
DR HSSP: P00763; IDPO.
DR MEROPS: S01.019; -.
DR MGD: MGI:1349451; Lrp4.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000024; Fz domain.
DR InterPro: IPR001272; LDL_receptor_A.
DR InterPro: IPR001254; Ser_protease_Try.
DR InterPro: IPR001190; Srcr_receptor.
DR Pfam: PF01392; Fz; 2.
DR Pfam: PF00057; ldl_recept_a; 6.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00261; LDLRECEPTOR.
DR SMART: SM00065; FRI; 2.
DR SMART: SM00192; LDLA; 3.
DR SMART: SM00202; SR; 1.
DR SMART: SM00020; Tryp_Spc; 1.
DR PROSITE: PS50038; FZ; 2.
DR PROSITE: PS01209; LDLRA_1; 6.
DR PROSITE: PS50068; LDLRA_2; 7.
DR PROSITE: PS00420; SRCR_1; FALSE_NEG.
DR PROSITE: PS50287; SRCR_2; 1.
DR PROSITE: PS50240; TRYPsin_DOM; 1.
DR PROSITE: PS00134; TRYPsin_HIS; FALSE_NEG.
DR PROSITE: PS00135; TRYPsin_SER; 1.
KW Hydrolyase; Serine protease; Transmembrane; Signal-anchor;
  Glycoprotein; Repeat.
FT DOMAIN 1 112
FT TRANSMEM 113 133
FT
FT DOMAIN 134 1113
FT DOMAIN 201 327
FT DOMAIN 336 372
FT DOMAIN 373 408
FT DOMAIN 409 445
FT DOMAIN 446 483
FT DOMAIN 518 641
FT DOMAIN 647 682
FT DOMAIN 683 721
FT DOMAIN 722 757

```

```

FT DOMAIN 758 853 SRCR.
FT ACT_SITE 869 1113 SERINE PROTEASE.
FT ACT_SITE 910 910 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 959 959 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 1052 1052 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 857 979 BY SIMILARITY.
FT DISULFID 895 911 BY SIMILARITY.
FT DISULFID 1022 1037 BY SIMILARITY.
FT DISULFID 1048 1077 BY SIMILARITY.
FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 202 202 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 298 298 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 317 317 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 373 373 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 411 411 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 444 444 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 519 519 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 635 635 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 719 719 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 765 765 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 828 828 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 970 970 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1089 1089 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1113 AA; 122984 MW; B845B2C5F20DD8EC CRC64;

```

```

Query Match 22.2%; Score 53.5; DB 1; Length 1113;
Best Local Similarity 36.0%; Pred. No. 45;
Matches 9; Conservative 5; Mismatches 8; Indels 3; Gaps 1;

```

```

Oy 8 WPMWRKHEAPEAE--PIMILKKW 29
Db 880 WPMWCSLQSEPSGICGVLIARKW 904

```

```

RESULT 15
ID FD6C_SOYBN STANDARD; PRT; 424 AA.
AC P48628;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Omega-6 fatty acid desaturase, chloroplast precursor (EC 1.14.19.-).
OS Glycine max (Soybean)
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=94345008; PubMed=8066133;
RA Hilt W.D., Carlson T.J., Booth J.R. Jr., Kinney A.J., Stecca K.L.,
RA Yadev N.S.;
RT "Cloning of a higher-plant plastid omega-6 fatty acid desaturase cDNA
RT and its expression in a cyanobacterium.";
RL Plant Physiol. 105:635-641(1994).
CC -1- FUNCTION: CHLOROPLAST OMEGA-6 FATTY ACID DESATURASE INTRODUCES
CC THE SECOND DOUBLEBOND IN THE BIOSYNTHESIS OF 16:3 AND 18:3 FATTY
CC ACIDS, IMPORTANT CONSTITUENTS OF PLANT MEMBRANES. IT IS THOUGHT
CC TO USE FERREDOXIN AS AN ELECTRON DONOR AND TO ACT ON FATTY ACIDS
CC ESTERIFIED TO GALACTOLIPIDS, SULFOLIPIDS AND PHOSPHATIDYLGLYCEROL.
CC -1- PATHWAY: Polyunsaturated fatty acid biosynthesis.
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST, MEMBRANE-BOUND (PROBABLE).
CC -1- DOMAIN: THE HISTIDINE BOX DOMAINS MAY CONTAIN THE ACTIVE SITE
CC AND/OR BE INVOLVED IN METAL ION BINDING.
CC -1- SIMILARITY: BELONGS TO THE FATTY ACID DESATURASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L29215; AAA50158.1;
CC PIR: T07742; T07742.
CC InterPro: IPR005804; FA_desat_fam.
CC Pfam: PF00487; FA_desaturase_1.
CC DR ProDom: PD001081; FA_desat_fam; 2.
CC Oxidoreductase; Fatty acid biosynthesis; Chloroplast; Membrane;
CC Transl. peptide.
CC TRANSIT 1 63 CHLOROPLAST (BY SIMILARITY).
CC CHAIN 64 424 OMEGA-6 FATTY ACID DESATURASE.
CC FT DOMAIN 165 169 HISTIDINE BOX-1.
CC FT DOMAIN 201 205 HISTIDINE BOX-2.
CC FT DOMAIN 361 365 HISTIDINE BOX-3.
CC SEQUENCE 424 AA; 49641 MW; 4F3DF52D4B1A2009 CRC64;

```

```

Query Match 22.0%; Score 53; DB 1; Length 424;
Best Local Similarity 34.4%; Pred. No. 20;
Matches 11; Conservative 7; Mismatches 12; Indels 2; Gaps 2;

```

```

Oy 5 WPMWRKHEAPEAEPIIMILKKWPMW-WR 35
Db 192 YPEPWRKHDRH-HAKTNMLREDTAMHPWVK 222

```

```

Search completed: July 29, 2003, 16:32:15
Job time : 14.6849 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2003, 16:28:05 ; Search time 60.8219 Seconds
(without alignments)
156.982 Million cell updates/sec

Title: SEQ35-27-35
Perfect score: 241
Sequence: 1 ILKKMPWMPWRKHEAPEEPIMLKKMPWMPWRK 37

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.todent:*
12: sp.virus:*
13: sp.vertibrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70.5	29.3	1245	3	Q9Y7V5
2	67.5	28.0	746	12	Q9YH31
3	67	27.8	723	12	Q9DU04
4	64	26.6	175	12	Q91RD8
5	63	26.1	49	12	Q9DT80
6	63	26.1	192	16	Q928B7
7	63	26.1	748	12	Q9D781
8	63	26.1	750	12	Q91D04
9	62	25.7	367	11	Q63778
10	61	25.3	92	12	Q8V7E2
11	61	25.3	250	17	Q8TWG0
12	61	25.3	284	10	Q94C18
13	61	25.3	734	12	Q8V711
14	61	25.3	735	12	Q9DUC9
15	61	25.3	985	15	Q98414
16	60	24.9	147	11	Q61427

17	60	24.9	242	17	Q8TVM8	Q8TVM8 methanopyru
18	60	24.9	363	16	Q8FHX9	Q8FHX9 escherichia
19	60	24.9	381	16	Q9A7E1	Q9A7E1 caulobacter
20	60	24.9	766	12	Q91CV5	Q91CV5 tt virus. o
21	59.5	24.7	114	16	Q9X8C2	Q9X8C2 streptomyce
22	59.5	24.7	2292	12	Q66765	Q66765 encephalomy
23	59	24.5	95	10	Q91QNO	Q91QNO arabidopsis
24	59	24.5	485	16	P72844	P72844 synechocyst
25	59	24.5	739	12	Q99A03	Q99A03 tt virus. o
26	58.5	24.3	521	10	Q94EP3	Q94EP3 oryza sativ
27	58.5	24.3	1567	2	Q9ADW1	Q9ADW1 polyangium
28	58	24.1	107	16	Q9XAE4	Q9XAE4 streptomyce
29	58	24.1	513	11	Q63289	Q63289 rattus norv
30	58	24.1	646	11	Q63779	Q63779 rattus norv
31	58	24.1	755	11	Q8CCF8	Q8CCF8 mus musculu
32	58	24.1	879	11	Q8V799	Q8V799 rattus norv
33	58	24.1	1300	11	P97692	P97692 rattus norv
34	57.5	23.9	117	15	Q9YR99	Q9YR99 human immun
35	57.5	23.9	341	13	Q90644	Q90644 gallus gall
36	57.5	23.9	376	3	Q94516	Q94516 schizosacch
37	57.5	23.9	444	4	Q9HC40	Q9HC40 homo sapien
38	57.5	23.9	772	4	Q9BXV6	Q9BXV6 homo sapien
39	57.5	23.9	1193	11	Q8BTG4	Q8BTG4 mus musculu
40	57.5	23.9	1194	11	Q8C0H4	Q8C0H4 mus musculu
41	57.5	23.9	1194	11	Q8BM10	Q8BM10 mus musculu
42	57	23.7	84	10	Q8H2U8	Q8H2U8 oryza sativ
43	57	23.7	252	15	Q91U37	Q91U37 human immun
44	57	23.7	426	12	Q99A07	Q99A07 tt virus. o
45	57	23.7	970	11	Q88821	Q88821 mus musculu

ALIGNMENTS

RESULT 1

ID Q9Y7V5 PRELIMINARY: PRT: 1245 AA.
AC Q9Y7V5; 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Candidiospore surface protein.
GN CMPI.
OS Trichoderma harzianum.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; mitosporic Hypocreales; Trichoderma.
OX NCBI_TaxID=5344;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 32173;
RX MEDLINE=99343881; PubMed=10413618;
RA Puyesky M., Benhamou N., Ponce Noyola P., Bauw G., Ziv T.,
RA Van Montagu M., Herrera Estrella A., Horvitz B.A.;
RT "Developmental regulation of cmpl, a gene encoding a multidomain
RT Candidiospore surface protein of Trichoderma.";
RL Fungal Genet. Biol. 27:88-99(1999).
DR EMBL: AJ133651; CAB40845.1;
DR HSSP: P01180; INPO.
DR InterPro: IPR006188; Claudin_sup.
DR InterPro: IPR001673; S_mold_repeat.
DR Prodom: PD006869; S_mold_repeat; 2.
DR PROSITE: PS01346; CLAUDIN: 1.
SQ SEQUENCE 1245 AA; 135624 MW; 3249C749AFA0CDF8 CRC64;

Query Match 29.3%; Score 70.5; DB 3; Length 1245;
Best local Similarity 29.0%; Pred. No. 2.3;
Matches 9; Conservative 2; Mismatches 9; Indels 11; Gaps 1;

QY 4 KMPWMPWRKHEAPEEPIMLKKMPWMPWRK 34
DB 1185 RMQWMSWPRRGRC-----CWQWMSW 1204

Query	Match	Similarity	Score	67.5	DB	12	Length	746
Best	Local	Similarity	33.3%	Pred	No.	3		
Matches	11	Conservative	2	Mismatches	5	Indels	15	Gaps
Q7	5	WFWWRKKHEAEPEAFIMILKKWFWWRKK	37					
Db	3	WGWWRRR-----RWPARRRRR	20					

Query match	27.88; Score 67; DB 12; Length 723;
-------------	-------------------------------------

Query Match	26.6%	Score 64	DB 12	Length 175
Best Local Similarity	33.3%	Pred. No. 2.1		
Matches	11	Conservative	2	Mismatches 4; Indels 16; Gaps 2.
Oy	5	WPPWPWRKRHEAPEAEPI	MLKKRPPWPPWRKK	37
	1	1	1	1
	1	1	1	1
Db	3	WSWP-WRRRR-----	-WPPRRR	19

Query Match	26.1%;	Score 63;	DB 12;	Length 49;
Best Local Similarity	30.3%;	Pred. No. 0.82;		
Matches 10;	Conservative 3;	Mismatches 4;	Indels 16;	Gaps 2

RESULT 6
Q9Z8B7

AD	092887	PRELIMINARY;	PRT;	192 AA.
DT	01-MAY-1999	(TREMBlrel. 10, Created)		
DT	01-MAY-1999	(TREMBlrel. 10, Last sequence update)		
DT	01-DEC-2001	(TREMBlrel. 19, Last annotation update)		
DE	CI277	SIMILARITY (CP00426 protein).		
GN	CP00426	OR CP00426 OR CP0327.		
OC	Chlamydia pneumoniae	(Chlamydophila pneumoniae).		
OC	Bacteria; Chlamydiae;	Chlamydiales; Chlamydiaceae; Chlamydophila.		
OX	NCBI_TaxID=83558;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-CWL029;			
RX	MEDLINE=99206606; PubMed=10192388;			
RA	Kalman S., Mitchell W., Marathe R., Lammell C., Fan J., Hyman R.W.,			
RA	Olinger L., Grimwood J., Davis R.W., Stephens R.S.;			
RT	"Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";			
RL	Nat. Genet. 21:385-389(1999).			
RL	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-AR39;			
RX	MEDLINE=20150255; PubMed=10684935;			
RA	Reed T.D., Brumham R.C., Shen C., Gill S.R., Heidelberg J.F.,			
RA	White O., Hickey E.K., Peterson J., Ullrich T., Berry K., Baas S.,			
RA	Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,			
RA	Gwynn M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,			
RT	Eisen J., Fraser C.M.;			
RT	"Genome sequences of Chlamydia trachomatis MOpn and Chlamydia			
RL	pneumoniae AR39.";			
RL	Nucleic Acids Res. 28:1397-1406(2000).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-J138;			
RX	MEDLINE=2030349; PubMed=10871362;			
RA	Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,			
RA	Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;			
RT	"Comparison of whole genome sequences of Chlamydia pneumoniae J138			
RL	from Japan and CWL029 from USA.";			
RL	Nucleic Acids Res. 28:2311-2314(2000).			
DR	EMBL; AE001625; AAD18570.1;			
DR	EMBL; AE002194; AAF8182.1;			
DR	EMBL; AP002546; BAA98634.1;			
DR	TIGR; CP0327;			
KW	Complete proteome.			
SO	SEQUENCE 192 AA; 21464 MW; 9638C329AEDDB76F CRC64;			
Query Match	26.1%;	Score 63;	DB 16;	Length 192;
Best Local Similarity	42.3%;	Pred. No. 3.1;		
Matches 11;	Conservative 4;	Mismatches 11;	Indels 0;	Gaps 0;
QY	5 WPMWPKRRKHEAPEADPEADIMILKKRP 30			
DB	138 WPMPLPKRQIEKLPKGEICFLSNYP 163			
RESULT 7				
Q9D781	PRELIMINARY;	PRT;	748 AA.	
AC	Q9D781;			
DT	01-MAR-2001	(TREMBlrel. 16, Created)		
DT	01-MAR-2001	(TREMBlrel. 16, Last sequence update)		
DT	01-OCT-2002	(TREMBlrel. 22, Last annotation update)		
DE	ORE1.			
OS	TT virus.			
OC	Viruses; ssDNA viruses; unclassified ssDNA viruses.			
OX	NCBI_TaxID=68887;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-TYM9;			
RX	MEDLINE=20568739; PubMed=11181348;			
RA	Okamoto H., Nishizawa T., Tawara A., Takahashi M., Kishimoto J.,			
RA	Sai T., Sugai Y.;			
RT	"TT virus mRNAs detected in the bone marrow cells from an infected			

```

RT individual":
RT Biochem. Biophys. Res. Commun. 279:700-707(2000).
DR EMBL: AB050448; BAB19928.1; -.
DR InterPro: IPR004219; TTVirus_Unk.
DR Pfam: PF02956; TT_ORF1; 1.
SQ SEQUENCE 748 AA; 88552 MW; D650CB2CA5CE26f CRC64;

Query Match
Best Local Similarity 30.3%; Pred. No. 12;
Matches 10; Conservative 3; Mismatches 4; Indels 16; Gaps 2;

OY 5 WPMWPMRRKHEAPEEPIMLKKWPMWPMRRK 37
DB ||| |::: ||||:
3 WTW-WQRRRR-----WPMRRR 19

RESULT 8
O91D04
AC O91D04 PRELIMINARY; PRT; 750 AA.
ID O91D04
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DR 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE ORF1.
OS TT virus.
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_TaxID=68887;
RN [1]
RP SEQUENCE FROM N.A.
RX Okamoto H., Nishizawa T., Takahashi M., Asabe S., Tsuda F.,
RA Yoshikawa A.;
RT "Heterogeneous distribution of TT virus of distinct genotypes in
RT multiple tissues from infected humans.";
RL Virology 288:358-368(2001).
DR EMBL: AB060592; BAB69900.1; -.
DR InterPro: IPR004219; TTVirus_Unk.
DR Pfam: PF02956; TT_ORF1; 1.
SQ SEQUENCE 750 AA; 89223 MW; 616EC86DC3469091 CRC64;

Query Match
Best Local Similarity 26.1%; Score 63; DB 12; Length 750;
Matches 10; Conservative 3; Mismatches 4; Indels 16; Gaps 2;

OY 5 WPMWPMRRKHEAPEEPIMLKKWPMWPMRRK 37
DB ||| |::: ||||:
3 WTW-WQRRRR-----WPMRRR 19

RESULT 9
O63778
ID O63778 PRELIMINARY; PRT; 367 AA.
AC O63778
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DR 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical 43.7 kDa protein.
DE Rattus norvegicus (Rat).
OC Eumariota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=87064324; PubMed=3023845;
RA D'Ambrosio E., Waltzkin S.D., Witney F.R., Salamee A., Furano A.V.;
RT "Structure of the highly repeated, long interspersed DNA family (LINE
RT or L1rn) of the rat.";
RL MOL. Cell. Biol. 6:411-424(1986).
DR EMBL: M13100; AAA66046.1; -.
DR InterPro: IPR000566; Lipoclin_cytraBP.
DR PROSITE: PS00213; LIPOCALIN; 1.
KW Hypothetical protein.

```

SO SEQUENCE 367 AA; 43685 MW; 465449CC185EC3B CRC64;

Query Match 25.7%; Score 62; DB 11; Length 367;
Best Local Similarity 39.4%; Pred. No. 7.7;
Matches 13; Conservative 4; Mismatches 12; Indels 4; Gaps 2;

OY 1 ILKKPMPWR---RKHEAPEAPIMILK-KW 29
DB 25 IFSKWCWFNMRATCRRMOIDPSLSPCTKLKSKW 57

RESULT 10

Q8V7E2

ID 08V7E2

PRELIMINARY;

PRT;

92 AA.

AC 08V7E2;

DT 01-MAR-2002 (TREMBlrel. 20, Created)

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)

DE 08F1 (Fragment).

OS TT virus.

OC Viruses; ssDNA viruses; unclassified ssDNA viruses.

OX NCBI_TaxID=68887;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=21844401; PubMed=11855633;

RA Peng Y.H., Nishizawa T., Takahashi M., Ishikawa T., Yoshikawa A.,

Okamoto H.;

"Analysis of the complete genomes of thirteen TT virus variants

classified into the fourth and fifth genetic groups, isolated from

Arch. Virol. 147:21-41(2002).

DR EMBL: AB064515; BAB79374.1;

DR InterPro: IPR004219; TTVirus_Unk.

DR Pfam: PF02956; TTV_ORF1; 1.

FT NON_TER

SO SEQUENCE 92 AA; 12429 MW; 188D883D05A7B09A CRC64;

Query Match

Best Local Similarity 32.3%;

Score 61; DB 12; Length 92;

Pred. No. 2.6;

Matches 10; Conservative 3; Mismatches 2; Indels 16; Gaps 1;

OY 7 WMPWRKHEAPEAPIMILKMPWMPWRK 37
DB 3 WMPWRKHEAPEAPIMILKMPWMPWRK 17

RESULT 11

O8TWG0

ID 08TWG0;

PRELIMINARY;

PRT;

250 AA.

AC 08TWG0;

DT 01-JUN-2002 (TREMBlrel. 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)

DE MAD-dependent protein deacetylase, SIR2 family.

GN SIR2 OR MK1075.

OS Methanopyrus kandleri.

OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;

OC Methanopyrus

OX NCBI_TaxID=2320;

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN=AV19 / DSM 6324 / JCM 9639;

RA MEDLINE=21927647; PubMed=11930014;

RA Slesarev A.I., Mezheva K.V., Makarova K.S., Polushin N.N.,

Shcherbina O.V., Shakhova V.V., Belova G.I., Aravind L.,

Malikh A.G., Koonin E.V., Kozayvkin S.A., Wolf Y.I., Stetter K.O.,

"The complete genome of hyperthermophile Methanopyrus kandleri AV19

RT and monophyly of archaeal methanogens";

Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).

DR EMBL: AE010396; AAM02288.1;

DR InterPro: IPR003000; SIR2.

DR Pfam: PF02146; SIR2; 1.

DR PROSITE: P550305; SIRTUN. 1.

KW Complete proteome.

SO SEQUENCE 250 AA; 27799 MW; B051994FE5B2AE05 CRC64;

Query Match 25.3%; Score 61; DB 17; Length 250;
Best Local Similarity 50.0%; Pred. No. 7;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

OY 3 KKWMPWRKHEAPEAP 22
DB 60 KWEMWLMRRRKIAEAPNP 79

RESULT 12

O94C18

ID 094C18

PRELIMINARY;

PRT;

284 AA.

AC 094C18;

DT 01-DEC-2001 (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE Glycine-rich protein LegRpl.

OS Lycopersicon esculentum (Tomato).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; Lamiales; Solanales; Solanaceae; Solanum.

OX NCBI_TaxID=4081;

RN [1]

RP SEQUENCE FROM N.A.

RA Uthappa M., Muchhal U.S., Baldwin J.C., Raghothama K.G.;

"LegRpl: A new member of glycine-rich proteins from tomato

RT (Lycopersicon esculentum)";

PL Physiol. Plantarum 0:0-0(2001).

DR EMBL: AY026037; AAK08984.1;

DR InterPro: IPR002952; Eggshell.

DR PRINTS: PR01228; EGGSHELL.

SO SEQUENCE 284 AA; 23434 MW; E81A84C247CB9ED8 CRC64;

Query Match

Best Local Similarity 31.2%;

Score 61; DB 10; Length 284;

Pred. No. 8;

Matches 10; Conservative 0; Mismatches 4; Indels 18; Gaps 2;

OY 5 WMPWRKHEAPEAPIMILKMPWMPWRK 36
DB 9 WMPWRKHEAPEAPIMILKMPWMPWRK 22

RESULT 13

O8V7I1

ID 08V7I1

PRELIMINARY;

PRT;

734 AA.

AC 08V7I1;

DT 01-MAR-2002 (TREMBlrel. 20, Created)

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)

DE ORF1.

OS TT virus.

OC Viruses; ssDNA viruses; unclassified ssDNA viruses.

OC NCBI_TaxID=68887;

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN=CT43F;

RA Okamoto H.;

Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

RC STRAIN=CT43F;

RA MEDLINE=21844401; PubMed=11855633;

RA Peng Y.H., Nishizawa T., Takahashi M., Ishikawa T., Yoshikawa A.,

Okamoto H.;

"Analysis of the complete genomes of thirteen TT virus variants

classified into the fourth and fifth genetic groups, isolated from

Arch. Virol. 147:21-41(2002).

DR EMBL: AB064598; BAB79322.1;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2003, 16:28:25 ; Search time 24.1644 Seconds
(without alignments)
143.272 Million cell updates/sec

Title: SEQ36-27-35
Perfect score: 236
Sequence: 1 ILRPMWPMRRKHEAPEEPIMLKKMPWPMRRK 36

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	33.1	144	1 JCI1222	indolicidin precursor
2	63	26.7	192	2 H86543	hypothetical prote
3	63	26.7	192	2 D12081	conserved hypotnet
4	61.5	26.1	2290	1 GNNYE	genome polypotein
5	60	25.4	381	2 B87470	hypothetical prote
6	59.5	25.2	114	2 T36208	hypothetical prote
7	59	25.0	95	2 E86447	protein F5D14.5 [i
8	59	25.0	451	2 S30401	hypothetical prote
9	59	25.0	485	2 S74708	hypothetical prote
10	58	24.6	711	2 C40046	antibiotic transpo
11	57.5	24.4	376	2 T40591	hypothetical prote
12	55.5	23.5	982	2 VCIJVS	env polypotein pr
13	55.5	23.5	983	1 E45390	env polypotein pr
14	55	23.3	352	2 S77448	hypothetical prote
15	55	23.3	452	2 T28094	hypothetical prote
16	55	23.3	715	2 B70741	probable moey prot
17	55	23.3	2292	2 S35961	capsid polypotein
18	54.5	23.1	134	2 E72532	hypothetical prote
19	54.5	23.1	128	2 AD3326	chloramphenicol O-
20	54.5	23.1	276	2 AH0244	probable esterase
21	54	22.9	144	2 S35331	apidaecin 22 precu
22	54	22.9	187	2 AB1647	hypothetical prote
23	54	22.9	424	2 T07742	omega-6 desaturase
24	54	22.9	469	2 B70607	probable cys prot
25	54	22.9	473	2 C86949	probable cysteinyl
26	54	22.9	1173	1 VCIHHC	E2 glycoprotein pr
27	53.5	22.7	538	2 B84759	hypothetical prote
28	53.5	22.7	990	1 G46335	env polypotein pr
29	53.5	22.7	1113	2 JEO315	low-density lipopr

30	53	22.5	107	2 T35634	hypothetical prote
31	53	22.5	314	2 S43916	hypothetical prote
32	53	22.5	406	2 H69143	coenzyme F420-redu
33	53	22.5	480	2 JC7552	Shb-like adapter p
34	53	22.5	691	2 D71450	hypothetical prote
35	53	22.5	949	2 E75352	glycine cleavage s
36	53	22.5	2292	1 GNNYED	genome polypotein
37	53	22.5	2292	1 GNNYED	genome polypotein
38	53	22.5	2292	1 S55401	capsid polypotein
39	52.5	22.2	55	2 T11026	H+-transporting tw
40	52.5	22.2	295	2 G90934	probable excinucle
41	52.5	22.2	295	2 C85783	probable excinucle
42	52.5	22.2	295	2 B64933	hypothetical prote
43	52.5	22.2	302	2 AD0709	conserved hypotnet
44	52.5	22.2	443	2 T08136	probable omega-6 d
45	52.5	22.2	448	2 D85362	hypothetical prote

ALIGNMENTS

RESULT 1

JCI1222
Indolicidin precursor - bovine
N:Alternate names: antimicrobial peptide
C:Species: Bos primigenius taurus (cattle)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: JCI1222; A42387; S25664
R:del Sal, G.; Storici, P.; Schneider, C.; Romeo, D.; Zanetti, M.
Biochem. Biophys. Res. Commun. 187, 467-472, 1992
A:Title: cDNA cloning of the neutrophil bactericidal peptide indolicidin.
A:Reference number: JCI1222; MUID:92392368; PMID:1520337
A:Accession: JCI1222
A:Molecule type: mRNA
A:Residues: 1-144 <SAL>
A:Cross-references: EMBL:X67340; NID:9462; PIDN:CAA47755.1; PID:9463
A:Experimental source: Bone marrow
R:Seidman, M.E.; Novotny, M.J.; Morris, W.L.; Tang, Y.Q.; Smith, W.; Cullor, J.S.
J. Biol. Chem. 267, 4292-4295, 1992
A:Title: Indolicidin, a novel bactericidal tridecapeptide amide from neutrophils.
A:Reference number: A42387; MUID:92165771; PMID:1537821
A:Accession: A42387
A:Molecule type: protein
A:Residues: 131-143 <SEL>
A:Experimental source: neutrophils
A>Note: Sequence extracted from NCBI backbone (NCBI:83840)
C:Superfamily: cathelin; cystatin homology
C:Keywords: amidated carboxyl end
F:1-29/Domain: signal sequence #status predicted <SIG>
F:32-129/Domain: cystatin homology <CYS>
F:30-130/Domain: propeptide #status predicted <PRO>
F:131-143/Product: indolicidin #status experimental <MAT>
F:143/Modified site: amidated carboxyl end (Arg) (amide in mature form from following

Query Match 33.1%; Score 78; DB 1; Length 144;
Best Local Similarity 55.6%; Pred. No. 0.0097;
Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 18 EAEPIMILKKMPWPMRR 35
DB 126 ELQSVILPMKWPMWPMRR 143

RESULT 2
H86543
hypothetical protein CPJ0426 [imported] - Chlamydomonas pneumoniae (strain J138)
C:Species: Chlamydomonas pneumoniae, Chlamydia pneumoniae
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C:Accession: H86543
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362

A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-114
A:Cross-references: EMBL:AL049763; PIDN:CAB42078.1; GSPDB:GN00070; SCOEDB:SCE36.09
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SCE36.09

Query Match 25.2%; Score 59.5; DB 2; Length 114;
Best Local Similarity 50.0%; Pred. No. 1.5;
Matches 10; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

OY 17 PEAEPIMLKKR-PMPMR 35
DB 93 PETAPADARRRPRMPMR 112

RESULT 7

E86447
protein F5D14.5 [Imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: E86447
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Huizart, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Maiti, R.; Matzali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: AB6141; MUID:21016719; PMID:11130712
A:Accession: E86447
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-95 <STO>
A:Cross-references: GB:AE005172; NID:g8920603; PIDN:AAF01325.1; GSPDB:GN00141
C:Genetics:
A:Gene: F5D14.5
A:Map position: 1

Query Match 25.0%; Score 59; DB 2; Length 95;
Best Local Similarity 21.3%; Pred. No. 1.4;
Matches 10; Conservative 5; Mismatches 4; Indels 28; Gaps 3;

OY 4 WP-----WMPRRKHEAPEEPIMLK---KMPMPW 33
DB 46 WPVVVVAVGCGGRMMWMM-----PVLVITDVGGSEWMMW 81

RESULT 8

S30401
hypothetical protein 2 - Streptomyces clavuligerus plasmid pSCL
C:Species: Streptomyces clavuligerus
C:Date: 30-Apr-1998 #sequence_revision 08-May-1998 #text_change 22-Oct-1999
C:Accession: S30401
R:Wu, X.; Roy, K.L.
J. Bacteriol. 175, 37-52, 1993
A:Title: Complete nucleotide sequence of a linear plasmid from Streptomyces clavuligerus
A:Reference number: S30400; MUID:93106972; PMID:8416908
A:Accession: S30401
A:Molecule type: DNA
A:Residues: 1-451 <WU>
A:Cross-references: EMBL:X54107; NID:948758; PIDN:CAA38041.1; PID:g561632
C:Genetics:
A:Genome: plasmid pSCL
A:Start codon: GTG

Query Match 25.0%; Score 59; DB 2; Length 451;
Best Local Similarity 25.0%; Pred. No. 7.3;
Matches 15; Conservative 3; Mismatches 4; Indels 38; Gaps 4;

OY 2 LRP-----WMPRRKHEAPEEPIMLKWP-----WMPW 33
DB 362 LRMPFEGFGLSPATGMAEELRLMDWLMR-----GPRSE-----RMPAOFCEGTPW 411

RESULT 9

S74708
hypothetical protein slr1306 - Synechocystis sp. (strain PCC 6803)
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C:Accession: S74708
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, O.; K. Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S74708
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-485 <KAN>
A:Cross-references: EMBL:D90901; GB:AB001339; NID:g1651897; PIDN:BAAL6859.1; PID:dl01
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 25.0%; Score 59; DB 2; Length 485;
Best Local Similarity 29.3%; Pred. No. 7.9;
Matches 12; Conservative 6; Mismatches 5; Indels 18; Gaps 2;

OY 5 PW-----NMPRRKHEAPEEPIIM---ILKK 27
DB 41 PWDQGLMALSGLVIMWRNRHNRPRQKQMLPREVLQK 81

RESULT 10

C40046
antibiotic transport-associated protein actII-3 - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 22-Oct-1999
C:Accession: C40046
R:Fernandez-Moreno, M.A.; Caballero, J.L.; Hopwood, D.A.; Maltipartida, F.
Cell 66, 769-780, 1991
A:Title: The act cluster contains regulatory and antibiotic export genes, direct targ
A:Reference number: A40046; MUID:91347376; PMID:1878971
A:Accession: C40046
A:Molecule type: DNA
A:Residues: 1-711 <FER>
A:Cross-references: GB:M64683; NID:g153143; PIDN:AAA26691.1; PID:g153146

Query Match 24.6%; Score 58; DB 2; Length 711;
Best Local Similarity 56.2%; Pred. No. 16;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 3 RMPMPWRKHEAPEE 18
DB 335 RMPMPWRKHEAPEE 350

RESULT 11

T40591
hypothetical protein SPBC646.15c - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 04-Mar-2000
C:Accession: T40591
R:Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, January 1999
A:Reference number: 221938
A:Accession: T40591
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-376 <SEB>
A:Cross-references: EMBL:AL035216; PIDN:CAA22819.1; GSPDB:GN00067; SPDB:SPBC646.15c

A:Experimental source: strain 972h-; cosmid c646
 C:Genetics:
 A:Gene: SPBC646.15c
 A:Map position: 2
 A:Insertions: 49/1.126/2; 312/2; 350/1
 C:Superfamily: Schizosaccharomyces pombe hypothetical protein SPBC646.15c

Query Match 24.4%; Score 57.5; DB 2; Length 376;
 Best local similarity 37.0%; Pred. No. 9.3;
 Matches 10; Conservative 3; Mismatches 7; Indels 7; Gaps 1;

QY 7 WPMRRKHEAPEEPIMLKKMPW 33
 DB 236 WPMRRKOKSS-----LKVPMGFW 255

RESULT 12

VCLJVS
 env polypeptide precursor - Maedi/Visna virus (strain 1514)

N:Alternate names: coat polypeptide
 C:Species: Maedi/Visna virus
 A:Note: host Homo sapiens (man)
 C:Date: 28-Feb-1986 #sequence_revision 28-Feb-1986 #text_change 16-Feb-1997
 C:Accession: A03977
 R:Sonigo, P.; Alizon, M.; Staskus, K.; Klatzmann, D.; Cole, S.; Danos, O.; Retzel, E.; T Cell 42: 369-382, 1985
 A:Title: Nucleotide sequence of the visna lentivirus: relationship to the AIDS virus.
 A:Reference number: A90869; MID:85254938; PMID:2410140
 A:Accession: A03977
 A:Molecule type: DNA
 A:Residues: 1-982 <SON>
 C:Genetics:
 A:Gene: env
 C:Superfamily: visna lentivirus type E retrovirus env polypeptide
 C:Keywords: capsid protein; coat protein; glycoprotein; polypeptide; transmembrane prote
 F:1-100/Domain: signal sequence #status predicted <SIG>
 F:1-101/Domain: signal sequence #status predicted <EXT>
 F:101-656/Product: exterior membrane glycoprotein #status predicted <TM>
 F:657-982/Product: transmembrane glycoprotein #status predicted <TM>
 F:140,161,206,258,298,364,370,381,387,403,414,435,439,470,475,481,491,501,515,527,537,54
 F:697,764,771,787,821/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 23.5%; Score 55.5; DB 1; Length 982;
 Best local similarity 34.3%; Pred. No. 45;
 Matches 12; Conservative 1; Mismatches 17; Indels 5; Gaps 1;

QY 1 ILRMPW-----PMRRKHEAPEEPIMLKKMPW 30
 DB 911 IMRATWMAKTSPMRHNRMPTLPLIVTQM 945

RESULT 13

E45390
 env polypeptide precursor - Maedi/Visna virus (strain KV1772) (provirus)

N:Alternate names: coat polypeptide
 N:Contains: exterior membrane glycoprotein; transmembrane glycoprotein
 C:Species: Maedi/Visna virus
 A:Note: host Homo sapiens (man)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
 C:Accession: E45390
 R:Andersson, O.S.; Elser, J.E.; Tobin, G.J.; Greenwood, J.D.; Gonda, M.A.; Georgsson, G.
 J.W.; Petursson, G.
 Virology 193, 89-105, 1993
 A:Title: Nucleotide sequence and biological properties of a pathogenic proviral molecule
 A:Reference number: A45390; MID:93174981; PMID:8382414
 A:Accession: E45390
 A:Molecule type: DNA
 A:Residues: 1-983 <NND>
 A:Cross-references: GB:S55323; NID:q265825; PIDN:AA825463.1; PID:q265830
 C:Genetics:
 A:Gene: env
 C:Superfamily: visna lentivirus type E retrovirus env polypeptide
 C:Keywords: capsid protein; coat protein; glycoprotein; polypeptide; transmembrane prote
 F:1-100/Domain: signal sequence #status predicted <SIG>

F:101-656/Product: exterior membrane glycoprotein #status predicted <EXT>
 F:657-983/Product: transmembrane glycoprotein #status predicted <TM>
 F:836-852/Domain: transmembrane #status predicted <TM>
 F:140,161,206,258,298,364,381,387,403,414,435,439,470,475,481,491,501,515,527,537,542

Query Match 23.5%; Score 55.5; DB 1; Length 983;
 Best local similarity 34.3%; Pred. No. 45;
 Matches 12; Conservative 1; Mismatches 17; Indels 5; Gaps 1;

QY 1 ILRMPW-----PMRRKHEAPEEPIMLKKMPW 30
 DB 912 IMRATWMAKTSPMRHNRMPTLPLIVTQM 946

RESULT 14

S77448
 hypothetical protein sl11080 - Synechocystis sp. (strain PCC 6803)

C:Species: Synechocystis sp.
 A:Variety: PCC 6803
 C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
 C:Accession: S77448
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, O.; K. Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas DNA Res. 3, 109-136, 1996
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis.
 A:Reference number: S74322; MID:97061201; PMID:8905331
 A:Accession: S77448
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-352 <KAN>
 A:Cross-references: EMBL:D90905; GB:AB001339; NID:g1652360; PIDN:BA17295.1; PID:d101
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 23.3%; Score 55; DB 2; Length 352;
 Best local similarity 52.6%; Pred. No. 18;
 Matches 10; Conservative 1; Mismatches 6; Indels 2; Gaps 2;

QY 17 PEAPIMI-LKKMP-WMPW 33
 DB 47 PEGPIVIGYSNMAGWMPW 65

RESULT 15

T28094
 hypothetical protein ZK899.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
 C:Accession: T28094
 R:Kerhaw, J.
 submitted to the EMBL Data Library, September 1994
 A:Reference number: Z20468
 A:Accession: T28094
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-452 <WIL>
 A:Cross-references: EMBL:Z37140; PIDN:CA85502.1; GSPDB:GN00028; CESP:ZK899.2
 A:Experimental source: clone ZK899
 C:Genetics:
 A:Gene: CESP:ZK899.2
 A:Map position: X
 A:Insertions: 34/3; 143/2; 227/2; 262/3; 380/3
 C:Superfamily: Caenorhabditis elegans hypothetical protein ZK899.2

Query Match 23.3%; Score 55; DB 2; Length 452;
 Best local similarity 32.1%; Pred. No. 23;
 Matches 9; Conservative 3; Mismatches 8; Indels 8; Gaps 2;

QY 6 WPMRRKHEAPEEPIMLKKMPW 33
 DB 175 WMTW---HDTDPN-----IFDRMNVPM 194

Tue Jul 29 17:06:39 2003

seq36-27-35.rpx

Page 5

Search completed: July 29, 2003, 16:35:21
Job time : 25.1644 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2003, 16:19:05 ; Search time 13.3151 Seconds

(without alignments)
127.146 Million cell updates/sec

Title: SEQ36-27-35
Perfect score: 236
Sequence: 1 ILRPMWPMWRKHEAPEADPIMILKKWPMWRK 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database : Swissprot_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	33.1	144	1	INDC_BOVIN
2	61.5	26.1	2290	1	POLG_EMCV
3	58	24.6	711	1	MLLA_STRCO
4	55.5	23.5	55	1	ARP8_ANAPL
5	55.5	23.5	982	1	ENV_VILV
6	55.5	23.5	983	1	ENV_VILV
7	55.5	23.5	991	1	ENV_VILV
8	55	23.3	715	1	YD5_MYCTU
9	54.5	23.1	942	1	ENV_CAEVG
10	54	22.9	144	1	AP22_APIME
11	54	22.9	424	1	FD6C_SOYBN
12	54	22.9	469	1	SYCL_MYCTU
13	54	22.9	473	1	SYCL_MYCTU
14	54	22.9	1173	1	VG12_CVH22
15	53.5	22.7	295	1	CHO_ECOLI
16	53.5	22.7	990	1	ENV_OMAVS
17	53.5	22.7	1042	1	CORI_HUMAN
18	53.5	22.7	1113	1	CORI_MOUSE
19	53	22.5	314	1	YMA3_BACST
20	53	22.5	2292	1	POLG_EMCV
21	53	22.5	2292	1	POLG_EMCV
22	52.5	22.2	257	1	ARP8_AYTAM
23	52.5	22.2	257	1	EA34_ADEA0
24	52.5	22.2	293	1	CHO_SALT1
25	52.5	22.2	293	1	CHO_SALT1
26	52.5	22.2	295	1	CHO_SALT1
27	52.5	22.2	295	1	CHO_SALT1
28	52.5	22.2	295	1	CHO_SALT1
29	52.5	22.2	443	1	FD6C_BRANA
30	52.5	22.2	448	1	FD6C_BRANA
31	52.5	22.2	691	1	YHOG_ECOLI
32	52.5	22.2	824	1	TGM1_RAT
33	52.5	22.2	989	1	ENV_VILV1

ALIGNMENTS

RESULT 1	ID	INDC_BOVIN	STANDARD:	PRT:	144 AA.
AC	P33046:				
DT	01-OCT-1993 (Rel. 27, Created)				
DT	01-OCT-1993 (Rel. 27, Last sequence update)				
DE	15-SEP-2003 (Rel. 42, Last annotation update)				
DE	Indolicidin precursor.				
OS	Bos taurus (Bovine).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
OC	Bovidae; Bovinae; Bos.				
OC	NCB1_TaxID=9913;				
OC	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=bone marrow;				
RC	TISSUE=bone marrow;				
RX	MEDLINE=9292368; PubMed=1520337;				
RA	del Sal G., Storici P., Schneider C., Romeo D., Zanetti M.;				
RT	"CDNA cloning of the neutrophil bactericidal peptide indolicidin.";				
RL	Biochem. Biophys. Res. Commun. 187:467-472(1992).				
RN	[2]				
RP	SEQUENCE OF 131-143.				
RC	TISSUE=Neutrophils;				
RX	MEDLINE=92165771; PubMed=1537821;				
RA	Selsted M.E., Novotny M.J., Morris W.L., Tang Y.-Q., Smith W.;				
RT	Cullor J.S.;				
RT	"Indolicidin, a novel bactericidal tridecapeptide amide from				
RT	neutrophils.";				
RL	J. Biol. Chem. 267:4292-4295(1992).				
CC	- FUNCTION: POTENT MICROBICIDAL ACTIVITY, ACTIVE AGAINST				
CC	STAPHYLOCOCCUS AUREUS AND ESCHERICHIA COLI.				
CC	- TISSUE SPECIFICITY: LARGE GRANULES OF NEUTROPHILS.				
CC	- PTM: ELASTASE MIGHT BE RESPONSIBLE FOR ITS MATURATION.				
CC	- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
CC	EMBL: X67340; CAA47755.1; ..				
DR	PIR: JCI222; JCI222.				
DR	PDB: 1G89; 17-JAN-01.				
DR	PDB: 1G8C; 17-JAN-01.				
DR	PDB: 1HRI; 31-DEC-02.				
DR	InterPro: IPR001894; Cathelicidin.				
DR	Pfam: PF00666; Cathelicidins; 1.				
DR	ProDom: PD001838; Cathelicidins; 1.				
DR	PROSITE: PS00946; CATHELICIDIN_1; 1.				
DR	PROSITE: PS00947; CATHELICIDIN_2; 1.				
KW	Antibiotic; Antidation; Signal; Pyroglutamate carboxylic acid;				
KW	3D-structure.				
FT	SIGNAL	1	29	POTENTIAL.	

P12722 avian infec
P11223 avian infec
P12650 avian infec
P12651 avian infec
P05135 avian infec
P02405 bacterioph
Q44241 anabaena sp
P17109 e menaquin
O11013 mycobacteri
Q09677 schizosach
P75771 escherichia
P03369 human immun

```

DR Pfam: PF00910: RNA_helicase: 1.
KW Polypeptin; Coat protein; Core protein; Transferrin;
RN directed RNA polymerase; Hydrolase; Thiol protease; Myristate.
FT PROPE 1 67 LEADER PEPTIDE.
FT CHAIN 68 136 COAT PROTEIN VP4 (RHO).
FT CHAIN 137 391 COAT PROTEIN VP2 (BETA).
FT CHAIN 392 622 COAT PROTEIN VP3 (GAMMA).
FT CHAIN 623 910 COAT PROTEIN VP1 (ALPHA).
FT CHAIN 911 1056 CORE PROTEIN P2A (G).
FT CHAIN 1057 1192 CORE PROTEIN P2B (I).
FT CHAIN 1193 1517 CORE PROTEIN P2C (F).
FT CHAIN 1518 1605 CORE PROTEIN P3A.
FT CHAIN 1606 1625 GENOME-LINKED PROTEIN VPG (H).
FT CHAIN 1626 1830 PICOORAIN 3C (P22).
FT CHAIN 1831 2290 MYRISTATE (BY SIMILARITY).
FT LIPID 68 68 MYRISTATE (POTENTIAL).
FT ACT_SITE 1784 1784 PROTEASE (POTENTIAL).
FT ACT_SITE 1802 1802 PROTEASE (POTENTIAL).
SQ SEQUENCE 2290 AA: 255756 MW: 268C81BB7CFC68CB5 CRC64;

Query Match 26.1%; Score 61.5; DB 1; Length 2290;
Best Local Similarity: 27.3%; Pred. No. 9.2;
Matches 12; Conservative 6; Mismatches 9; Indels 17; Gaps 1

OY 3 RWFWMPWRKRRHEAPEAEPI-----MILKWP 29
Db 965 RAFPWNPWKNTGYALVAREPCRVMTDLYKKVRFLPLVQKEMP 1008
:::|::|

RESULT 3
MMLA_STRCO
ID MMLA_STRCO STANDARD; PRT; 711 AA.
AC Q53902;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative membrane protein actII-3.
GN ActII-3 OR SC05084 OR SCHAC28G1.10.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91347376; PubMed=1878971;
RA Fernandez-Moreno M.A., Caballero J.L., Hopwood D.A., Malpartida F.;
RA "The act cluster contains regulatory and antibiotic export genes,
RA direct targets for translational control by the hlda tRNA gene of
RA streptomycetes."
RA Cell 66:769-780(1991).
RL [2]
RE SEQUENCE FROM N.A.
RX STRAIN=A3(2) / ML45;
RX MEDLINE=2196410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabschawski E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RA "Complete genome sequence of the model actinomycete Streptomyces
RA coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE MMLA FAMILY
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```



```

FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 414 414 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 433 433 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 439 439 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 470 470 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 475 475 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 491 491 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 501 501 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 515 515 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 527 527 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 543 543 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 568 568 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 697 697 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 764 764 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 771 771 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 787 787 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 821 821 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT VARIANT 12 12 T -> M (IN REF. 2)
FT VARIANT 118 118 S -> N (IN REF. 2)
FT VARIANT 283 283 K -> R (IN REF. 2)
FT VARIANT 641 641 ER -> AQ (IN REF. 2)
FT VARIANT 645 645 R -> K (IN REF. 2)
SQ SEQUENCE 982 AA; 113978 MW; 7D78BAE6E22BF53F CRC64;

```

```

Query Match 23.5%; Score 55.5; DB 1; Length 982;
Best Local Similarity 34.3%; Pred. No. 22;
Matches 12; Conservative 1; Mismatches 17; Indels 5; Gaps 1;
QY 1 ILRPMW----PWRKHEAPEEPIMLKKMPW 30
Db 911 IWRATWMAWKTSPMRHNRMTPTLPLVIMQW 945

```

```

ENV_VILV2 STANDARD; PRT; 983 AA.
ID ENV_VILV2
AC P35954;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE ENV polyprotein precursor (Coat polyprotein).
GN ENV.
OS Vigna lentivirus (strain KVI1772).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=36374;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93174981; PubMed=8382414;
RA Anderson O.S., Elser J.E., Tobin G.J., Greenwood J.D., Gonda M.A.,
RA Georgsson G., Andresdottir V., Benediktsson E., Carlsson H.M.,
RA Meenylae E.O., Rafnar B., Palsson P.A., Casey J.W., Petursson G.;
RT "Nucleotide sequence and biological properties of a pathogenic
RT proviral molecular clone of neuroinfectious virus."
RL Virology 193:89-105(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: S55323; AAB25463.1; -
CC EMBL: I06906; AAA48362.1; -
CC PDB: 1JRK; 25-JUL-01.
CC InterPro: IPR000328; Env.GP41.
CC Pfam: PF00517; GP41; 1.
CC GlycoProtet: Coat protein; Polypeptide; Transmembrane; 3D-structure.
KW PEPTIDE
FT 1 100 LEADER PEPTIDE.

```

```

FT CHAIN 101 656 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT TRANSMEMBRANE GLYCOPROTEIN.
FT POTENTIAL.
FT CARBOHYD 836 832 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 161 161 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 206 206 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 258 258 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 364 364 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 381 381 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 414 414 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 435 435 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 439 439 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 470 470 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 475 475 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 491 491 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 501 501 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 515 515 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 527 527 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 543 543 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 568 568 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 697 697 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 765 765 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 772 772 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 788 788 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 822 822 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CONFLICT 900 900 R -> H (IN REF. 1; AAA48362).
SQ SEQUENCE 983 AA; 113939 MW; E03BC254EA78268 CRC64;

```

```

Query Match 23.5%; Score 55.5; DB 1; Length 983;
Best Local Similarity 34.3%; Pred. No. 22;
Matches 12; Conservative 1; Mismatches 17; Indels 5; Gaps 1;
QY 1 ILRPMW----PWRKHEAPEEPIMLKKMPW 30
Db 912 IWRATWMAWKTSPMRHNRMTPTLPLVIMQW 946

```

```

ENV_VILV2 STANDARD; PRT; 991 AA.
ID ENV_VILV2
AC P23423;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ENV polyprotein precursor (Coat polyprotein).
GN ENV.
OS Vigna lentivirus (strain 1514 / clone IWI-1KS2).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=11744;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91134986; PubMed=1847257;
RA Staakus K.A., Retzel E.F., Lewis E.D., Wietegrefe S.W., Silsby J.L.,
RA Cyr S., Rank J.M., Haase A.T., Fast D., Geisler P.T., Hardy J.T.,
RA Kong S.H., Cook R., Lahli C.J., Neufeld T.P., Porter T.E., Shoop E.,
RA Zachow K.R.;
RT "Isolation of replication-competent molecular clones of Vigna virus."
RL Virology 181:228-240(1991).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```



```

DR EMBL: M60610; AAA17532.1; ALT_INIT.
DR InterPro: IPR00328; Env_GP41.
DR Pfam: PF00517; GP41. 1.
KW Glycoprotein; Coat protein; Polypeptide; Transmembrane.
FT PEPTIDE 1 100
FT CHAIN 101 664 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT TRANSMEM 665 991 TRANSMEMBRANE GLYCOPROTEIN.
FT CARBOHYD 844 860 POTENTIAL.
FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 161 161 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 206 206 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 258 258 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 298 298 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 364 364 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 381 381 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 435 435 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 439 439 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 470 470 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 475 475 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 491 491 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 501 501 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 515 515 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 527 527 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 543 543 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 568 568 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 705 705 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 773 773 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 780 780 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 796 796 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 830 830 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 17 17 E -> D (IN REF. 1; AAA17532).
SQ SEQUENCE 991 AA; 115016 MW; D8920FF4A8BA55A7 CRC64;

Query Match 23.5%; Score 55.5; DB 1; Length 991;
Best Local Similarity 34.3%; Pred. No. 22;
Matches 12; Conservative 1; Mismatches 17; Indels 5; Gaps 1;

OY 1 ILRPMW-----PWRKHEAPEDEPIMLKMPW 30
DB 920 IWRATWMAWKTSPPMRHNRTPMYITLLPLIVIMQW 954

RESULT 8
YD55_MYCTU STANDARD: PRT; 715 AA.
AC 011025;
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein RV1355c.
GN RV1355C OR MT1398 OR MTCY02B10.19C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RC MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
RA Horsby T., Jagsels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulten J.E., Taylor K., Whitehead S., Barrall B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the

```

```

RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Biswal W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RT Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch)
CC -----
DR EMBL: Z75555; CAAG9988.1;
DR EMBL: AE007012; AAK4561.1; ALT_INIT.
DR PIR: B70741; B70741.
DR TIGR: MT1398;
DR Tuberculist: RV1355c;
DR InterPro: IPR00594; Thif_domain.
DR Pfam: PF00899; Thif; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 715 AA; 78181 MW; 455495248A56041C CRC64;

Query Match 23.3%; Score 55; DB 1; Length 715;
Best Local Similarity 38.1%; Pred. No. 19;
Matches 8; Conservative 7; Mismatches 4; Indels 2; Gaps 1;

OY 17 PEAPDIMI--LKKMPMPWRR 35
DB 53 PQPDDLLAEAKRMAYYPWRR 73

RESULT 9
ENV_CAEGV STANDARD: PRT; 942 AA.
AC P31627;
DT 01-JUL-1993 (Rel. 26, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Env polypeptide precursor (Coat polypeptide) [Contains: Surface
DE protein; Transmembrane protein].
GN ENV.
OS Caprine arthritis encephalitis virus (strain G63) (CAEV).
OC Viruses; Retrovirus; Retroviridae; Lentivirus.
OX NCBI_TaxID=11662;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 81-95.
RX MEDLINE=92015464; PubMed=1656067;
RA Knowles D.P. Jr., Cheevers W.P., McGuire T.C., Brassfield A.L.,
RA Harwood W.G., Stem T.A.;
RT "Structure and genetic variability of envelope glycoproteins of two
RT antigenic variants of caprine arthritis encephalitis lentivirus."
RL J. Virol. 65:5744-5750(1991).
RN [2]
RP REVISIONS.
RA Knowles D.P.;
RP Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch)
CC -----

```

$$\mathbb{R}^N \quad [2]$$

RT amplification of the insect antibacterial response.";

CC -I- PATHWAY: Polyunsaturated fatty acid biosynthesis.

CC -|- PATHWAY: Polyunsaturated fatty acid biosynthesis.

```

CC -I- DOMAIN: THE HISTIDINE BOX DOMAINS MAY CONTAIN THE ACTIVE SITE
CC AND/ OR BE INVOLVED IN METAL ION BINDING.
CC -I- SIMILARITY: BELONGS TO THE FATY ACID DESATURASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L29215; AAA50158.1;
CC PIR: T07742; T07742.
CC InterPro: IPR005804; FA_desat.fam.
CC Pfam: PF00487; FA_desaturase.1.
CC ProDom: PD001081; FA_desat.fam.2.
CC Oxidoreductase; Fatty acid biosynthesis; Chloroplast; Membrane;
CC Transf.
CC TRANSIT 1 63 CHLOROPLAST (BY SIMILARITY).
CC CHAIN 64 424 OMEGA-6 FATTY ACID DESATURASE.
CC DOMAIN 165 169 HISTIDINE BOX-1.
CC DOMAIN 201 205 HISTIDINE BOX-2.
CC DOMAIN 361 365 HISTIDINE BOX-3.
CC SEQUENCE 424 AA; 49641 MW; 4F3DF52D4B1A2009 CRC64;

Query Match 22.9%; Score 54; DB 1; Length 424;
Best Local Similarity 35.3%; Pred. No. 15;
Matches 12; Conservative 7; Mismatches 13; Indels 2; Gaps 2;

Qy 2 LRPMPWPRRRKHEAPEEPIILKKMPWP-WR 34
Db 190 LIPYEPYRFRKHDRH-HAKTNMLREDTAMHPYWK 222

RESULT 12
SYCL_MYCTU STANDARD: PRT; 469 AA.
AC P96862;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE CysteinyI-tRNA synthetase 1 (EC 6.1.1.16) (Cysteine--tRNA ligase 1)
DE (CYRS 1).
GN CYSS1 OR CYSS OR RV3580C OR MT3686 OR MTCY06G11.27C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinomycetales;
OC Corynebacteriineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96895987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares R.,
RA Sultun J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RA complete genome sequence."
RA Nature 393:537-544 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CD1151 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A.L., Utterback T., Weidman J., Khouri L., Gill J., Mikula A.,
RA Bishai W.;
RA "Whole genome comparison of Mycobacterium tuberculosis clinical and

```

```

RT Laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -I- CATALYTIC ACTIVITY: ATP + L-Cysteine + tRNA(Cys) = AMP +
CC diphosphate + L-cysteinyl-tRNA(Cys).
CC -I- SUBUNIT: Monomer (By similarity).
CC -I- SUBCELLULAR LOCATION: Cytoplasmic.
CC -I- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC strong, to methionyl-tRNA synthetase.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z92774; CAB07154.1;
CC PIR: B70607; B70607.
CC TIGR: MT3686;
CC TubercuList: RV3580C;
CC HAMAP: MF_00041; -.
CC InterPro: IPR002308; Cys-tRNA-synt_1a.
CC InterPro: IPR001412; tRNA-synt_1.
CC Pfam: PF01406; tRNA-synt_1e; 1.
CC PRINTS: PR00983; TRNASYNTHCS.
CC TIGREMS: TIGR00435; cyss; 1.
CC PROSITE: PS00178; AA-tRNA-LIGASE_1; FALSE_NEG.
CC Aminoacyl-tRNA synthetase; Protein biosynthesis; ligase; ATP-binding;
CC Complete proteome.
CC SITE 35 45 "HIGH" REGION.
CC SITE 267 271 "KMSKS" REGION.
CC BINDING 270 270 ATP (BY SIMILARITY).
CC CONFLICT 457 457 D -> E (IN REF. 2).
CC SEQUENCE 469 AA; 51854 MW; 515F9D19482ADC0E CRC64;

Query Match 22.9%; Score 54; DB 1; Length 469;
Best Local Similarity 34.5%; Pred. No. 17;
Matches 10; Conservative 0; Mismatches 7; Indels 12; Gaps 1;

Qy 5 PMPWPRRRKHE-----APEEAP 21
Db 89 PMPWMAAHERAFATAYDALDVLPSAEP 117

RESULT 13
SYCL_MYCLE STANDARD: PRT; 473 AA.
AC P57990;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE CysteinyI-tRNA synthetase 1 (EC 6.1.1.16) (Cysteine--tRNA ligase 1)
DE (CYRS 1).
GN CYSS1 OR CYSS OR ML0323.
OS Mycobacterium lepreae.
OC Bacteria; Actinobacteria; Actinomycetales;
OC Corynebacteriineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RC MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duboy S., Fellwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;

```

```

RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
CC -1- CATALYTIC ACTIVITY: ATP + L-cysteine + tRNA(Cys) = AMP +
CC diphosphate + L-cysteinyl-tRNA(Cys).
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC Strong, to methionyl-tRNA synthetase.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AL583918; CAC29831.1; -.
CC PIR: C86949; C86949.
CC DR Leproma; ML0323; -.
CC DR HAMAP: MF_00041; -. 1.
CC DR InterPro: IPR002308; Cys-tRNA-synt_1a.
CC DR InterPro: IPR001412; tRNA-synt_1.
CC DR Pfam: PF01406; tRNA-synt_1e; 1.
CC DR PRINTS: PR00983; tRNAsyntHCys.
CC DR TIGRfam: TIGR00435; cysS; 1.
CC DR PROSITE: PS00178; AA-TRNA-LIGASE_I; FALSE_NEG.
CC KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC Complete proteome.
CC FT SITE 35 "HIGH" REGION.
CC FT SITE 267 "KMSKS" REGION.
CC FT BINDING 270 ATP (BY SIMILARITY).
CC FT SEQUENCE 473 AA; 52124 MW; 9PD6CF859C69316D CRC64;
SQ
Query Match 22.9%; Score 54; DB 1; Length 473;
Best Local Similarity 34.5%; Pred. No. 17;
Matches 10; Conservative 0; Mismatches 7; Indels 12; Gaps 1;
OY 5 PMWPRRKHE-----APEAPEP 21
DB 89 PMWMAATHERAFTNAIDALDYLPEPSAP 117
RESULT 14
VGL2_CVH22 STANDARD: PRT; 1173 AA.
AC P15423; P89342; P89343; P89344; Q66174; Q990M1; Q990M2; Q990M3;
AC Q990M4;
AC 01-APR-1990 (rel. 14, Created)
AC 01-APR-1990 (rel. 14, Last sequence update)
DT 15-SEP-2003 (rel. 42, Last annotation update)
DE E2 glycoprotein precursor (Spike glycoprotein) (Peplomer protein).
GN S.
OS Human coronavirus (strain 229E) (HCoV-229E).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=11137;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90264837; PubMed=2345367;
RA Raabe T., Schelle-Prinz B., Siddell S.G.;
RT "Nucleotide sequence of the gene encoding the spike glycoprotein of
RT human coronavirus HCV 229E.";
RL J. Gen. Virol. 71:1065-1073(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21262210; PubMed=11369870;
RA Thiel V., Herold J., Schelle B., Siddell S.G.;
RT "Infectious genome transcribed in vitro from a cDNA copy of the human
RT coronavirus RNA cloned in vaccinia virus.";
RL J. Gen. Virol. 82:1273-1281(2001).
RN [3]
RP SEQUENCE OF 98-1113 FROM N.A., AND VARIANTS.

```

```

RC STRAIN-Isolate RW Stock, Isolate P100E, Isolate P11A, and
RC Isolate P11B;
RA Bonavia A., Holmes K.V.;
RT "Viral and cellular changes in a human cell line persistently infected
RT with human coronavirus HCoV-229E.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 98-1113 FROM N.A., AND VARIANTS.
RC STRAIN-Isolate ATCC VR-74, Isolate A162, and Isolate LRI 281;
RX MEDLINE=99086140; PubMed=9870593;
RA Hays J.P., Myint S.H.;
RT "PCR sequencing of the spike genes of geographically and
RT chronologically distinct human coronaviruses 229E.";
RL J. Virol. Methods 75:179-193(1998).
RN [5]
RP SEQUENCE OF 1159-1173 FROM N.A.
RX MEDLINE=6936667; PubMed=2701946;
RA Raabe T., Siddell S.;
RT "Nucleotide sequence of the human coronavirus HCV 229E mRNA 4 and mRNA
RT 5 unique regions.";
RL Nucleic Acids Res. 17:6387-6387(1989).
RN [6]
RP INTERACTION WITH ANPEP.
RX MEDLINE=22440020; PubMed=12551991;
RA Bonavia A., Zelus B.D., Wentworth D.E., Talbot P.J., Holmes K.V.;
RT "Identification of a receptor-binding domain of the spike glycoprotein
RT of human coronavirus HCoV-229E.";
RL J. Virol. 77:2530-2538(2003).
RN [7]
RP INTERACTION WITH ANPEP.
RX MEDLINE=22521439; PubMed=12634402;
RA Breslin J.J., Mork T., Smith M.K., Vogel L.K., Hammla E.M.,
RA Bonavia A., Talbot P.J., Sjoestrom H., Noren O., Holmes K.V.;
RT "Human coronavirus 229E: Receptor binding domain and neutralization by
RT soluble receptor at 37 degrees C.";
RL J. Virol. 77:4435-4438(2003).
RN [8]
RP REVIEW.
RX MEDLINE=21109095; PubMed=11162792;
RA Gallagher T.M., Buchmeier M.J.;
RT "Coronavirus spike proteins in viral entry and pathogenesis.";
RL Virology 279:371-374(2001).
CC -1- FUNCTION: Structural protein that makes spikes at the surface of
CC the virus. Determines enteropathogenicity and virulence of the
CC virus. Initiates infection by specifically recognizing and binding
CC the human aminopeptidase ANPEP receptor. Its association with
CC ANPEP may lead to its conformational change that triggers fusion
CC between viral and host cellular membrane.
CC -1- SUBUNIT: Homotrimer. During virus morphogenesis, it is found in a
CC complex with M and HE proteins (By similarity). Interacts with
CC ANPEP.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- DOMAIN: The spike S1 domain displays the specificity for the host
CC receptor.
CC -1- DOMAIN: The leucine zipper-like heptad repeats may mediate the
CC fusion of viral and cellular membranes.
CC -1- POLYMORPHISM: The strong variation between the different
CC strains may affect the virulence of the virus.
CC -1- MISCELLANEOUS: In contrast to serogroup 2, E2 glycoprotein protein
CC from serogroup 1 is not cleaved.
CC -1- SIMILARITY: Contains 1 spike S1 domain.
CC -1- SIMILARITY: Contains 1 spike S2 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X16816; CAA34723.1; -.
CC DR EMBL: AF304460; AM648592.1; -.

```

Query Match	Best Local	Similarity	Score 54:	DB 1:	Length 1173:	
Matches	6:	Conservative	Pred. No. 40:	Mismatches	1: Indels	0: Gaps
QY	27	KMPWMPW 33				
DB	1113	KMPWMPW 1119				
RESULT 15						
CHO_ECOL6		STANDARD:				
AC	P59439:					
DT	15-SEP-2003 (Rel. 42, Created)					
DT	15-SEP-2003 (Rel. 42, Last sequence update)					
DT	15-SEP-2003 (Rel. 42, Last annotation update)					
DE	Excludenlease cho (EC 3.1.25.-) (Endonuclease cho) (uvrc homolog protein).					
GN	CHO OR C2140.					
OS	Escherichia coli O6.					
OC	Bacteria: Proteobacteria; Gammaproteobacteria; Enterobacteriales;					
OX	NCBI_TaxID=217992;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=O6:H1 / CFT073 / ATCC 700928;					
RX	MEDLINE=22388234; PubMed=12471157;					
RA	Weich R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,					
RA	Raske D., Buckles E.L., Liu S.-R., Boutin A., Hackett J., Stroud D.,					
RA	Mayhew G.F., Rose D.J., Zhou S.-R., Schwartz D.C., Pena N.T.,					
RA	Mobley H.L.T., Donnenberg M.S., Blattner F.R.;					
RT	"Extensive mosaic structure revealed by the complete genome sequence					
RT	of uropathogenic Escherichia coli.":					
RT	Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).					
RL						
CC	-I- FUNCTION: Incises the DNA at the 3' side of a lesion during					
CC	nucleotide excision repair. Incises the DNA farther away from the					
CC	lesion than uvrc. Not able to incise the 5' site of a lesion. When					
CC	a lesion remains because uvrc is not able to induce the 3'					
CC	incision, cho incises the DNA. Then uvrc makes the 5' incision.					
CC	The combined action of cho and uvrc broadens the substrate range					
CC	of nucleotide excision repair (By similarity).					
CC	-I- SIMILARITY: TO THE N-TERMINAL PART OF THE UVRC FAMILY.					
CC						
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -					
CC	the European Bioinformatics Institute. There are no restrictions on its					

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2003, 16:28:05 / Search time 59.1781 Seconds
(without alignments)
156.982 Million cell updates/sec

Title: SEQ36-27-35
Perfect score: 236
Sequence: 1 ILRWPMWPMRRKHEAPEADPIMILKKWPMWPMRRK 36

Scoring table: BLOSUM62
Gapop 10.0, Gapept 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

SPTREMBL_23:*
1: sp-archaea:*
2: sp-bacteria:*
3: sp-fungi:*
4: sp-human:*
5: sp-invertebrate:*
6: sp-mammal:*
7: sp-mhc:*
8: sp-organelle:*
9: sp-phage:*
10: sp-plant:*
11: sp-rodent:*
12: sp-virus:*
13: sp-vertebrate:*
14: sp-unclassified:*
15: sp-ryivrus:*
16: sp-bacteriap:*
17: sp-archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73.5	31.1	1245	3	Q917V5
2	68.5	29.0	746	12	Q9JH31
3	67	28.4	723	12	Q9JH31
4	65	27.5	175	12	Q9JH31
5	64	27.1	49	12	Q9JH31
6	64	27.1	748	12	Q9JH31
7	64	27.1	750	12	Q9JH31
8	63	26.7	147	11	Q61427
9	63	26.7	192	16	Q92887
10	62	26.3	735	12	Q9JH31
11	61.5	26.1	2292	12	Q66765
12	61	25.8	92	12	Q8V7E2
13	61	25.8	252	15	Q91T6
14	61	25.8	252	15	Q91U04
15	61	25.8	252	15	Q91T26
16	61	25.8	252	15	Q91T07

17	61	25.8	284	10	Q94C18	Q94C18 lycopersico
18	61	25.8	734	12	Q8V711	Q8V711 tt virus. o
19	61	25.8	766	12	Q91CY5	Q91CY5 tt virus. o
20	60	25.4	381	16	Q9A7E1	Q9A7E1 caulobacter
21	59.5	25.2	114	16	Q9X8C2	Q9X8C2 streptomyces
22	59	25.0	95	10	Q9LQ80	Q9LQ80 arabidopsis
23	59	25.0	102	16	Q8P429	Q8P429 xanthomonas
24	59	25.0	105	16	Q8P429	Q8P429 xanthomonas
25	59	25.0	252	15	Q91T09	Q91T09 human immun
26	59	25.0	451	2	Q05074	Q05074 streptomyces
27	59	25.0	485	16	P72844	P72844 synechocyst
28	59	25.0	739	12	Q99AQ3	Q99AQ3 tt virus. o
29	58.5	24.8	521	10	Q94EF3	Q94EF3 oryza sativ
30	58.5	24.8	985	15	Q98A14	Q98A14 ovine lentiv
31	58.5	24.8	1567	2	Q9ADM1	Q9ADM1 polyangium
32	58	24.6	250	17	Q8TW60	Q8TW60 methanopyru
33	58	24.6	252	15	Q91U37	Q91U37 human immun
34	58	24.6	252	15	Q91U06	Q91U06 human immun
35	58	24.6	426	12	Q99A07	Q99A07 tt virus. o
36	58	24.6	755	11	Q8C8F8	Q8C8F8 mus musculu
37	58	24.6	879	11	Q8V199	Q8V199 rattus norv
38	57.5	24.4	376	3	Q94516	Q94516 schizosach
39	57	24.2	252	15	Q91T05	Q91T05 human immun
40	57	24.2	445	10	Q8LQ06	Q8LQ06 oryza sativ
41	57	24.2	562	15	Q99B86	Q99B86 human immun
42	57	24.2	754	12	Q9JH33	Q9JH33 tt virus. o
43	57	24.2	754	12	Q8U2K7	Q8U2K7 sen virus.
44	57	24.2	970	11	Q88821	Q88821 mus musculu
45	57	24.2	971	11	Q70458	Q70458 mus musculu

ALIGNMENTS

RESULT 1
ID 0917V5 PRELIMINARY: PRT: 1245 AA.
AC 0917V5:

DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DR Candidospora surface protein.
GN CMPL.
OS Trichoderma harzianum.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; mitosporic Hypocreales; Trichoderma.
OX NCBI_TaxID=5544;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 32173;
RX MEDLINE=9343881; PubMed=10413618;
RA Puyesky M., Benhamou N., Ponce Noyola P., Bauw G., Ziv T.,
Van Montagu M., Herrera Estrella A., Horvitz B.A.;
RT "Developmental regulation of cmpl, a gene encoding a multidomain
Candidospora surface protein of Trichoderma.";
RL Fungal Genet. Biol. 27:88-99(1999).
DR EMBL: A133651; CAB40845.1; -;
DR HSSP: P01180; INPO.
DR InterPro: IPR006188; Claudin_sup.
DR InterPro: IPR001673; S.mold.repeat.
DR Prodom: PD006869; S.mold.repeat. 2.
DR PROSITE: PS01346; CLAUDIN: 1.
SQ SEQUENCE 1245 AA; 135824 MW; 3249C749AFA0CF8 CRC64;

Query Match 31.1%; Score 73.5; DB 3; Length 1245;
Best Local Similarity 32.3%; Pred. No. 0.95;
Matches 10; Conservative 1; Mismatches 9; Indels 11; Gaps 1;

QY 3 RMPWPMWPMRRKHEAPEADPIMILKKWPMWPMW 33
DB 1185 RMPWPMWPMRRKHEAPEADPIMILKKWPMWPMW 1204

20.46; score 61; DB 12; Length 723;

RESULT 6

0181


```

ID Q9DT81 PRELIMINARY; PRT: 748 AA.
AC Q9DT81;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, last annotation update)
DE ORF1.
RN
OS TT virus.
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_TaxID=68887;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=TYM9;
RX MEDLINE=20568739; PubMed=11118348;
RA Okamoto H., Nishizawa T., Tawara A., Takahashi M., Kishimoto J.,
RA Sai T., Sugai Y.;
RT "TT virus mRNAs detected in the bone marrow cells from an infected
RT individual."
RL Blochem. Biophys. Res. Commun. 279:700-707(2000).
DR EMBL; AB050448; BAB19928.1;
DR InterPro: IPR004219; TVVirus_Unk.
DR Pfam: PF02956; TT_ORF1.1.
SQ SEQUENCE 748 AA; 88552 MW; D65CGB2CA5CE26F CRC64;

Query Match 27.1%; Score 64; DB 12; Length 748;
Best Local Similarity 28.6%; Pred. No. 8.4;
Matches 10; Conservative 4; Mismatches 5; Indels 16; Gaps 2;

OY 2 LRMPMPWRKHEAPEAPIMILKKMPWPRRK 36
DB 1 MAMTW-MORRRR-----WPRRR 19
      ||| |||
      ||| |||

- RESULT 7
O91D04 PRELIMINARY; PRT: 750 AA.
ID Q91D04
AC Q91D04;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, last annotation update)
DE ORF1.
RN
OS TT virus.
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_TaxID=68887;
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=21488921; PubMed=11601907;
RA Okamoto H., Nishizawa T., Takahashi M., Asabe S., Tsuda F.,
RA Yoshikawa A.;
RT "Heterogeneous distribution of TT virus of distinct genotypes in
RT multiple tissues from infected humans."
RL Virology 288:358-368(2001).
DR EMBL; AB060592; BAB69900.1;
DR InterPro: IPR004219; TVVirus_Unk.
DR Pfam: PF02956; TT_ORF1.1.
SQ SEQUENCE 750 AA; 89223 MW; 616EC86DC3469091 CRC64;

Query Match 27.1%; Score 64; DB 12; Length 750;
Best Local Similarity 28.6%; Pred. No. 8.5;
Matches 10; Conservative 4; Mismatches 5; Indels 16; Gaps 2;

OY 2 LRMPMPWRKHEAPEAPIMILKKMPWPRRK 36
DB 1 MAMTW-MORRRR-----WPRRR 19
      ||| |||
      ||| |||

- RESULT 8
O61427 PRELIMINARY; PRT: 147 AA.
ID O61427
AC O61427;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, last annotation update)
DE Alpha 1 type I collagen (Fragment).

```

```

GN COL1A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Liver;
RX MEDLINE=94344105; PubMed=8065328;
RA Rhodes K., Rippe R.A., Umezawa A., Nehls M., Brenner D.A., Breindl M.;
RT "DNA methylation represses the murine alpha 1(I) collagen promoter by
RT an indirect mechanism."
RL Mol. Cell. Biol. 14:5950-5960(1994).
DR EMBL; X54876; CAA38657.1;
DR MGD; MGI:88467; COL1A1.
DR InterPro: IPR001007; VWF-C.
DR Pfam: PF00093; VWC; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01208; VWF-C; 1.
KW Collagen.
FT NON_TER
SQ SEQUENCE 147 AA; 16652 MW; 9263BF0A91B4307D CRC64;

Query Match 26.7%; Score 63; DB 11; Length 147;
Best Local Similarity 36.6%; Pred. No. 2.3;
Matches 15; Conservative 0; Mismatches 8; Indels 18; Gaps 4;

OY 3 RMPMPWRKHEAPEAPIMILKKW----PR----WPR 34
DB 116 RWP--PWTRW-----PWTAMTSWSSSPWAPRTPWPR 147
      ||| |||
      ||| |||

- RESULT 9
O928B7 PRELIMINARY; PRT: 192 AA.
ID O928B7
AC O928B7;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE CT27 SIMILARITY (CPJ0426 protein).
GN CPN0426 OR CPJ0426 OR CPJ327.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis."
RL Nat. Genet. 21:385-389(1999).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brumham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Ultebeck T., Berry K., Bass S.,
RA Liner K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin J., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RT pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shida T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA."
RL Nucleic Acids Res. 28:2311-2314(2000).

```

```

encephalomyocarditis virus RNA
RL J. Virol. 66:1602-1609(1992).
DR EMBL: M81861; AAA43037.1; -.
DR HSSP: P12296; IMC.
DR InterPro: IPR001676; Rbv.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF00073; rhv_3.
DR Pfam: PF00680; RNA_dep_RNA_pol_1.
DR Pfam: PF00910; RNA_helicase_1.
DR ProSite: PSS0507; RDRP_POSITIVE; 1.
DR ProSite: PSS0521; RDRP_VIRAL; 1.
FT CHAIN 68 137 PROTEIN 1A.
FT CHAIN 138 393 PROTEIN 1B.
FT CHAIN 394 624 PROTEIN 1C.
FT CHAIN 625 901 PROTEIN 1D.
FT CHAIN 902 1044 PROTEIN 2A.
FT CHAIN 1045 1194 PROTEIN 2B.
FT CHAIN 1195 1519 PROTEIN 2C.
FT CHAIN 1520 1627 PROTEIN 3AB.
FT CHAIN 1628 1832 PROTEIN 3C.
FT CHAIN 1833 2292 PROTEIN 3D.
SQ SEQUENCE 2292 AA; 255457 MW; 01C0537888CFC94 CRC64;

Query Match 26.1%; Score 61.5; DB 12; Length 2292;
Best Local Similarity 27.3%; Pred. No. 51;
Matches 12; Conservative 6; Mismatches 9; Indels 17; Gaps 1;

QY 3 RPPWPPRRKHEAPEEAPL-----MIKKWP 29
Db 967 RAPPWPKNTYQAVLRAPPCVMTDIYKRVPRFPLVQKWP 1010

RESULT 12
08V7E2 PRELIMINARY; PRT; 92 AA.
ID 08V7E2
AC 08V7E2;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE - ORF1 (Fragment).
OS TT virus.
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_TaxID=68887;
RN SEQUENCE FROM N.A.
RP MEDLINE=21844401; Pubmed=11855633;
RA Peng Y.H.; Nishizawa T., Takahashi M., Ishikawa T., Yoshikawa A.,
RA Okamoto H.;
RT "Analysis of the complete genomes of thirteen TT virus variants
RT classifiable into the fourth and fifth genetic groups, isolated from
RT viremic infants.";
RL Arch. Virol. 147:21-41(2002).
RL EMBL: AB064615; BAB79374.1; -.
DR InterPro: IPR004219; TTVirus-Unk.
DR Pfam: PF02956; TT-ORF1; 1.
FT NON_TER 92 92
SQ SSEQUENCE 92 AA; 12429 MW; 188D883D05A7B09A CRC64;

Query Match 25.8%; Score 61; DB 12; Length 92;
Best Local Similarity 32.3%; Pred. No. 2.6;
Matches 10; Conservative 3; Mismatches 2; Indels 16; Gaps 1;

QY 6 WPPRRKHEAPEEAPLMIKKWPPRRK 36
Db 3 WMYRRR-----PWRPWRK 17

RESULT 13
Q9ITX6 PRELIMINARY; PRT; 252 AA.
ID Q9ITX6
```

```

AC Q91RX6;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE V-1 reverse transcriptase (Fragment).
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97451;
RX MEDLINE=21102417; PubMed=11158089;
RA Servais J., Lambert C., Fontaine E., Plessier J.M., Robert I.,
RA Arendt V., Staub T., Schneider F., Hemmer R., Burtonboy G.,
RA Schmitt J.C.;
RT "Comparison of DNA sequencing and a line probe assay for detection of
RT human immunodeficiency virus Type 1 in patients failing highly active
RT antiretroviral therapy."
RL J. Clin. Microbiol. 39:454-459(2001).
DR EMBL: AJ401789; CAB86558.1; -.
DR InterPro: IPR000477; RVTse.
DR Pfam: PF00078; rvt; 1.
KW RNA-directed DNA polymerase; Transferase.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 252 AA; 29450 MW; 8E48D6280726D12 CRC64;

```

```

Query Match 25.8%; Score 61; DB 15; Length 252;
Best Local Similarity 27.3%; Pred. No. 6.8;
Matches 12; Conservative 8; Mismatches 10; Indels 14; Gaps 1;

```

```

OY 1 ILRPMWPMRRKHEAP-----EAEPIMLKKMPW 30
DB 209 LRLMGWYTPDKKHQKEPPLMGYELHPDKWTVPVLEPKDSW 252

```

RESULT*14

```

ID Q91U04 PRELIMINARY; PRT: 252 AA.
AC Q91U04;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE V-1 reverse transcriptase (Fragment).
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97001;
RX MEDLINE=21102417; PubMed=11158089;
RA Servais J., Lambert C., Fontaine E., Plessier J.M., Robert I.,
RA Arendt V., Staub T., Schneider F., Hemmer R., Burtonboy G.,
RA Schmitt J.C.;
RT "Comparison of DNA sequencing and a line probe assay for detection of
RT human immunodeficiency virus Type 1 in patients failing highly active
RT antiretroviral therapy."
RL J. Clin. Microbiol. 39:454-459(2001).
DR EMBL: AJ401761; CAB86530.1; -.
DR InterPro: IPR000477; RVTse.
DR Pfam: PF00078; rvt; 1.
KW RNA-directed DNA polymerase; Transferase.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 252 AA; 29441 MW; F2B8F8B9A8334D9C CRC64;

```

```

Query Match 25.8%; Score 61; DB 15; Length 252;
Best Local Similarity 27.3%; Pred. No. 6.8;
Matches 12; Conservative 8; Mismatches 10; Indels 14; Gaps 1;

```

```

OY 1 ILRPMWPMRRKHEAP-----EAEPIMLKKMPW 30
DB 209 LRLMGWYTPDKKHQKEPPLMGYELHPDKWTVPVLEPKDSW 252

```

```

RESULT 15
ID Q91RX6 PRELIMINARY; PRT: 252 AA.
AC Q91RX6;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE V-1 reverse transcriptase (Fragment).
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97095;
RX MEDLINE=21102417; PubMed=11158089;
RA Servais J., Lambert C., Fontaine E., Plessier J.M., Robert I.,
RA Arendt V., Staub T., Schneider F., Hemmer R., Burtonboy G.,
RA Schmitt J.C.;
RT "Comparison of DNA sequencing and a line probe assay for detection of
RT human immunodeficiency virus Type 1 in patients failing highly active
RT antiretroviral therapy."
RL J. Clin. Microbiol. 39:454-459(2001).
DR EMBL: AJ401769; CAB86538.1; -.
DR InterPro: IPR000477; RVTse.
DR Pfam: PF00078; rvt; 1.
KW RNA-directed DNA polymerase; Transferase.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 252 AA; 29346 MW; 7CB1C1ECD16750CF CRC64;

```

```

Query Match 25.8%; Score 61; DB 15; Length 252;
Best Local Similarity 27.3%; Pred. No. 6.8;
Matches 12; Conservative 8; Mismatches 10; Indels 14; Gaps 1;

```

```

OY 1 ILRPMWPMRRKHEAP-----EAEPIMLKKMPW 30
DB 209 LRLMGWYTPDKKHQKEPPLMGYELHPDKWTVPVLEPKDSW 252

```

```

Search completed: July 29, 2003, 16:34:25
Job time : 60.1781 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2003, 16:29:40 ; Search time 19.726 Seconds

(without alignments)
77.217 Million cell updates/sec

Title: SEQ36-27-35
Perfect score: 236
Sequence: 1 ILRWPMPWRRKHEAPEAPIMILKKMPWMPWRRK 36

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents-AA:
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/CTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	152.5	64.6	28	4	US-09-030-619-104
2	148.5	62.9	28	4	US-09-030-619-50
3	130.5	55.3	63	4	US-09-099-631A-12
4	130.5	55.3	63	4	US-09-416-481A-39
5	98.5	41.7	21	3	US-08-915-314-46
6	98.5	41.7	21	4	US-09-030-619-47
7	98.5	41.7	21	4	US-09-667-486-46
8	94.5	40.0	15	4	US-09-030-619-39
9	94	39.8	20	3	US-08-915-314-47
10	94	39.8	20	4	US-09-030-619-24
11	94	39.8	20	4	US-09-667-486-47
12	93.5	39.6	21	3	US-08-915-314-48
13	93.5	39.6	21	4	US-09-030-619-48
14	93.5	39.6	21	4	US-09-667-486-48
15	91	38.6	13	3	US-08-915-314-30
16	91	38.6	13	3	US-08-915-314-62
17	91	38.6	13	3	US-08-915-314-63
18	91	38.6	13	3	US-08-915-314-64
19	91	38.6	13	3	US-09-042-071-36
20	91	38.6	13	4	US-09-030-619-95
21	91	38.6	13	4	US-09-030-619-99
22	91	38.6	13	4	US-09-667-486-30
23	91	38.6	13	4	US-09-667-486-62
24	91	38.6	13	4	US-09-667-486-63
25	91	38.6	13	4	US-09-667-486-64
26	91	38.6	14	3	US-08-915-314-57
27	91	38.6	14	4	US-09-030-619-72

28	91	38.6	14	4	US-09-030-619-108	Sequence 108, App
29	91	38.6	14	4	US-09-667-486-57	Sequence 57, Appl
30	91	38.6	21	3	US-08-915-314-54	Sequence 54, Appl
31	91	38.6	21	4	US-09-030-619-69	Sequence 69, Appl
32	91	38.6	21	4	US-09-667-486-54	Sequence 54, Appl
33	87	36.9	12	3	US-08-915-314-52	Sequence 67, Appl
34	87	36.9	12	4	US-09-030-619-67	Sequence 52, Appl
35	87	36.9	12	4	US-09-667-486-52	Sequence 52, Appl
36	86	36.4	12	3	US-08-915-314-42	Sequence 42, Appl
37	86	36.4	12	3	US-08-915-314-74	Sequence 74, Appl
38	86	36.4	12	3	US-08-702-054B-5	Sequence 5, Appl1
39	86	36.4	12	4	US-09-030-619-23	Sequence 23, Appl
40	86	36.4	12	4	US-09-030-619-112	Sequence 112, Appl
41	86	36.4	12	4	US-09-667-486-42	Sequence 42, Appl
42	86	36.4	12	4	US-09-667-486-74	Sequence 74, Appl
43	86	36.4	13	3	US-08-915-314-51	Sequence 51, Appl
44	86	36.4	13	3	US-08-915-314-58	Sequence 58, Appl
45	86	36.4	13	3	US-08-702-054B-34	Sequence 34, Appl

ALIGNMENTS

```
RESULT 1
US-09-030-619-104
; Sequence 104, Application US/09030619B
; Patent No. 6503881
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; FILE REFERENCE: 660081.406
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 104
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-104

Query Match      64.6%  Score 152.5; DB 4; Length 28;
Best Local Similarity 63.9%; Pred. No. 5.2e-13;
Matches 23; Conservative 1; Mismatches 1; Indels 11; Gaps 1;

OY 1 ILRWPMPWRRKHEAPEAPIMILKKMPWMPWRRK 36
Db 1 ILRWPMPWRRK-----ILMKPMPWMPWRRK 25

RESULT 2
US-09-030-619-50
; Sequence 50, Application US/09030619B
; Patent No. 6503881
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; FILE REFERENCE: 660081.406
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 104
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cationic Peptide Analogue
```

```
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 50
LENGTH: 28
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-50

Query Match
Best Local Similarity 62.9%; Score 148.5; DB 4; Length 28;
Matches 24; Conservative 1; Mismatches 0; Indels 11; Gaps 2;

QY 1 ILRWPMPWRKHEAPEAPIMILKKWMPWRR 36
Db 1 ILRWPMPWRRK-----MIL-RWPMPWRRK 25

RESULT 3
US-09-099-631A-12
; Sequence 12, Application US/09099631A
; Patent No. 6444645
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; TITLE OF INVENTION: Crosslink-stabilized Indolicidin Analogs
; FILE REFERENCE: P-UC 3050
; CURRENT APPLICATION NUMBER: US/09/099,631A
; CURRENT FILING DATE: 1998-06-18
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 12
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-099-631A-12

Query Match
Best Local Similarity 55.3%; Score 130.5; DB 4; Length 63;
Matches 21; Conservative 1; Mismatches 5; Indels 7; Gaps 2;

QY 3 RWPMPWRKHEAPEAPIMILKKWMPWRR 35
Db 11 KWPMPWRRM-----ARIAMILPKWMPWRR 38

RESULT 4
US-09-416-481A-39
; Sequence 39, Application US/09416481A
; Patent No. 6524585
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; TITLE OF INVENTION: Indolicidin Analogs and Methods of Using Same
; FILE REFERENCE: P-UC 3794
; CURRENT APPLICATION NUMBER: US/09/416,481A
; CURRENT FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: US 09/076,227
; PRIOR FILING DATE: 1998-05-12
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 39
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: construct
```

```
US-09-416-481A-39

Query Match
Best Local Similarity 55.3%; Score 130.5; DB 4; Length 63;
Matches 21; Conservative 1; Mismatches 5; Indels 7; Gaps 2;

QY 3 RWPMPWRKHEAPEAPIMILKKWMPWRR 35
Db 11 KWPMPWRRM-----ARIAMILPKWMPWRR 38

RESULT 5
US-08-915-314-46
; Sequence 46, Application US/08915314
; Patent No. 6180604
; GENERAL INFORMATION:
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfle, Douglas
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
; NUMBER OF SEQUENCES: 90
; CORRESPONDENCE ADDRESS:
; ADDRESS: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,314
; FILING DATE: 20-AUG-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6180604tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 660081.405
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-915-314-46

Query Match
Best Local Similarity 41.7%; Score 98.5; DB 3; Length 21;
Matches 18; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

QY 1 ILRWPMPWRKHEAPEAPIMILKK 27
Db 1 ILRWPMPWRRK-----IMILKK 18

RESULT 6
US-09-030-619-47
; Sequence 47, Application US/09030619B
; Patent No. 6503881
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfle, Douglas
; APPLICANT: Fraser, Janet R.
```

APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
TITLE OF INVENTION: WITH ANTIBIOTICS
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 47
LENGTH: 21
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-47

Query Match 41.7%; Score 98.5; DB 4; Length 21;
Best Local Similarity 66.7%; Pred. No. 2.4e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

QY 1 ILRWPMPWRRKHEAPEAPIMILKK 27
1 ILRWPMPWRRK-----IMILKK 18

RESULT 7
US-09-667-486-46
Sequence 46, Application US/09667486
Patent No. 6538106
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
West, Michael H.P.
Krieger, Timothy J.
Taylor, Robert
Erfile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/667,486
FILING DATE: 22-Sep-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: No. 6538106tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 46:

US-09-667-486-46

Query Match 41.7%; Score 98.5; DB 4; Length 21;
Best Local Similarity 66.7%; Pred. No. 2.4e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

QY 1 ILRWPMPWRRKHEAPEAPIMILKK 27
1 ILRWPMPWRRK-----IMILKK 18

RESULT 8
US-09-030-619-39
Sequence 39, Application US/09030619B
Patent No. 6503881
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfile, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
TITLE OF INVENTION: WITH ANTIBIOTICS
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 39
LENGTH: 15
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-39

Query Match 40.0%; Score 94.5; DB 4; Length 15;
Best Local Similarity 41.7%; Pred. No. 5.4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 21; Gaps 1;

QY 1 ILRWPMPWRRKHEAPEAPIMILKKWMPWRRK 36
1 ILRWPW-----WMPWRRK 15

RESULT 9
US-08-915-314-47
Sequence 47, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/08/915,314
; FILING DATE: 20-AUG-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6180604tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 660081.405
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
US-08-915-314-47

Query Match          39.8%; Score 94; DB 3; Length 20;
Best Local Similarity 63.0%; Pred. No. 8.5e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 ILRWPMPWRRKHAEPEAPIMILK 27
Db 1 ILRWPMPWRRK-----MILK 17

RESULT 10
US-09-030-619-24
; Sequence 24, Application US/09030619B
; Patent No. 6503861
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfle, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 20
; TYPE: PPT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indollicidin Analogue
;
US-09-030-619-24

Query Match          39.8%; Score 94; DB 4; Length 20;
Best Local Similarity 63.0%; Pred. No. 8.5e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 ILRWPMPWRRKHAEPEAPIMILK 27
Db 1 ILRWPMPWRRK-----MILK 17

RESULT 11
US-09-667-486-47
; Sequence 47, Application US/09667486
; Patent No. 6538106
; GENERAL INFORMATION:
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfle, Douglas
```

```

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; INFECTIONS USING ANALOGUES OF INDOLICIDIN
; NUMBER OF SEQUENCES: 90
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/667,486
; FILING DATE: 22-Sep-2000
; CLASSIFICATION: <Unknown>
; PRIORITY INFORMATION:
; APPLICATION NUMBER: US/08/915,314
; FILING DATE: 20-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6538106tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 660081.405
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
;
INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
;
SEQUENCE DESCRIPTION: SEQ ID NO: 47:
;
US-09-667-486-47

Query Match          39.8%; Score 94; DB 4; Length 20;
Best Local Similarity 63.0%; Pred. No. 8.5e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 ILRWPMPWRRKHAEPEAPIMILK 27
Db 1 ILRWPMPWRRK-----MILK 17

RESULT 12
US-08-915-314-48
; Sequence 48, Application US/08915314
; Patent No. 6180604
; GENERAL INFORMATION:
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfle, Douglas
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
; NUMBER OF SEQUENCES: 90
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
CURRENT APPLICATION DATA:
```


APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tensburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-48

Query Match 39.6%; Score 93.5; DB 3; Length 21;
Best Local Similarity 63.0%; Pred. No. 1e-05;
Matches 17; Conservative 0; Mismatches 1; Indels 9; Gaps 1;

QY 1 ILRPMWPMRRKHEAPEAPIMILKK 27
Db 1 ILRPMWPMRRKD-----MILKK 18

RESULT 13
US-09-030-619-48
Sequence 48, Application US/09030619B
Patent No. 6503881
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfle, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 48
LENGTH: 21
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Indoliciidin Analogue
US-09-030-619-48

Query Match 39.6%; Score 93.5; DB 4; Length 21;
Best Local Similarity 63.0%; Pred. No. 1e-05;
Matches 17; Conservative 0; Mismatches 1; Indels 9; Gaps 1;

QY 1 ILRPMWPMRRKHEAPEAPIMILKK 27
Db 1 ILRPMWPMRRKD-----MILKK 18

RESULT 14
US-09-667-486-48
Sequence 48, Application US/09667486
Patent No. 6538106
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfle, Douglas

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/667,486
FILING DATE: 22-Sep-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: No. 6538106tensburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: <unknown>
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 48:
US-09-667-486-48

Query Match 39.6%; Score 93.5; DB 4; Length 21;
Best Local Similarity 63.0%; Pred. No. 1e-05;
Matches 17; Conservative 0; Mismatches 1; Indels 9; Gaps 1;

QY 1 ILRPMWPMRRKHEAPEAPIMILKK 27
Db 1 ILRPMWPMRRKD-----MILKK 18

RESULT 15
US-08-915-314-30
Sequence 30, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfle, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-30

Query Match 38.6%; Score 91; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 24 ILKKPWPWPWRRK 36
|||||
Db 1 ILKKPWPWPWRRK 13

Search completed: July 29, 2003, 16:36:07
Job time : 20.726 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2003, 16:34:31 ; Search time 33.0411 Seconds
(without alignments)
129.395 Million cell updates/sec

Title: SEQ36-27-35
Perfect score: 236
Sequence: 1 ILRMPWMPWRRKHEAPEAPRIMLKMPWMPWRRK 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published_Applications_MA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PC1_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PC1US_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	152.5	64.6	28	US-09-030-619-104	Sequence 104, App
2	148.5	62.9	28	US-09-030-619-50	Sequence 50, Appl
3	98.5	41.7	21	US-09-030-619-47	Sequence 47, Appl
4	94.5	40.0	15	US-09-030-619-39	Sequence 39, Appl
5	94	39.8	20	US-09-030-619-24	Sequence 24, Appl
6	93.5	39.6	21	US-09-030-619-48	Sequence 48, Appl
7	91	38.6	13	US-09-030-619-95	Sequence 95, Appl
8	91	38.6	13	US-09-030-619-99	Sequence 99, Appl
9	91	38.6	13	US-10-252-773-3	Sequence 3, Appl1
10	91	38.6	14	US-09-030-619-72	Sequence 72, Appl
11	91	38.6	14	US-09-030-619-108	Sequence 108, App
12	91	38.6	21	US-09-030-619-67	Sequence 67, Appl
13	87	36.9	12	US-09-030-619-67	Sequence 23, Appl
14	86	36.4	12	US-09-030-619-23	Sequence 112, App
15	86	36.4	12	US-09-030-619-112	Sequence 112, App

16	86	36.4	13	US-09-030-619-53	Sequence 53, Appl
17	86	36.4	13	US-09-030-619-107	Sequence 107, App
18	86	36.4	13	US-09-030-619-109	Sequence 109, App
19	86	36.4	14	US-09-030-619-54	Sequence 54, Appl
20	86	36.4	14	US-09-030-619-110	Sequence 110, App
21	86	36.4	20	US-09-030-619-51	Sequence 51, Appl
22	86	36.4	21	US-09-030-619-52	Sequence 52, Appl
23	85	36.0	12	US-09-030-619-73	Sequence 73, Appl
24	85	36.0	13	US-09-030-619-41	Sequence 41, Appl
25	85	36.0	13	US-09-030-619-103	Sequence 103, App
26	83	35.2	12	US-09-030-619-43	Sequence 43, Appl
27	83	35.2	12	US-09-030-619-44	Sequence 44, Appl
28	83	35.2	13	US-09-030-619-105	Sequence 105, App
29	83	35.2	13	US-09-030-619-106	Sequence 106, App
30	82	34.7	12	US-09-030-619-30	Sequence 30, Appl
31	82	34.7	12	US-09-030-619-111	Sequence 11, App
32	82	34.7	13	US-09-030-619-58	Sequence 58, Appl
33	82	34.7	13	US-09-030-619-59	Sequence 59, Appl
34	82	34.7	13	US-09-030-619-94	Sequence 94, Appl
35	82	34.7	13	US-09-030-619-102	Sequence 102, App
36	81	34.3	11	US-09-030-619-79	Sequence 79, Appl
37	81	34.3	11	US-09-030-619-114	Sequence 114, App
38	81	34.3	12	US-09-030-619-82	Sequence 82, Appl
39	81	34.3	12	US-09-030-619-91	Sequence 91, Appl
40	80	33.9	12	US-09-030-619-83	Sequence 83, Appl
41	80	33.9	12	US-09-030-619-89	Sequence 89, Appl
42	80	33.9	12	US-09-030-619-90	Sequence 90, Appl
43	78	33.1	11	US-09-030-619-78	Sequence 78, Appl
44	78	33.1	11	US-09-030-619-113	Sequence 113, App
45	78	33.1	12	US-09-030-619-31	Sequence 31, Appl

ALIGNMENTS

RESULT 1
US-09-030-619-104
Sequence 104, Application US/09030619B
Patent No. US20020035061A1
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 104
LENGTH: 28
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-104

Query Match 64.6%; Score 152.5; DB 9; Length 28;
Best Local Similarity 63.9%; Pred. No. 1.5e-10;
Matches 23; Conservative 1; Mismatches 1; Indels 11; Gaps 1;

OY 1 ILRMPWMPWRRKHEAPEAPRIMLKMPWMPWRRK 36
|||||
DB 1 ILRMPWMPWRRK-----ILRMPWMPWRRK 25
RESULT 2
US-09-030-619-50

```
; Sequence 50, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 50
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-50
```

```
Query Match 62.9%; Score 148.5; DB 9; Length 28;
Best Local Similarity 66.7%; Pred. No. 4.3e-10;
Matches 24; Conservative 1; Mismatches 0; Indels 11; Gaps 2;
```

```
QY 1 ILRWPMWPMWRKHAEPEEPIMLIKKPMWPMWRK 36
      |||||
DB 1 ILRWPMWPMWRK-----ML-RWPMWPMWRK 25
```

```
RESULT 3
US-09-030-619-47
; Sequence 47, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 47
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-47
```

```
Query Match 41.7%; Score 98.5; DB 9; Length 21;
Best Local Similarity 66.7%; Pred. No. 0.00011;
Matches 18; Conservative 0; Mismatches 0; Indels 9; Gaps 1;
```

```
QY 1 ILRWPMWPMWRKHAEPEEPIMLIKK 27
      |||||
DB 1 ILRWPMWPMWRK-----IMILKK 18
```

```
RESULT 4
US-09-030-619-39
; Sequence 39, Application US/09030619B
```

```
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-39
```

```
Query Match 40.0%; Score 94.5; DB 9; Length 15;
Best Local Similarity 41.7%; Pred. No. 0.00023;
Matches 15; Conservative 0; Mismatches 0; Indels 21; Gaps 1;
```

```
QY 1 ILRWPMWPMWRKHAEPEEPIMLIKKPMWPMWRK 36
      |||||
DB 1 ILRWPMW-----WPMWPMWRK 15
```

```
RESULT 5
US-09-030-619-24
; Sequence 24, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-24
```

```
Query Match 39.8%; Score 94; DB 9; Length 20;
Best Local Similarity 63.0%; Pred. No. 0.00033;
Matches 17; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
```

```
QY 1 ILRWPMWPMWRKHAEPEEPIMLIKK 27
      |||||
DB 1 ILRWPMWPMWRK-----MILKK 17
```

```
RESULT 6
US-09-030-619-48
; Sequence 48, Application US/09030619B
; Patent No. US20020035061A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erle, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTI-BIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030.619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 48
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin Analogue
; US-09-030-619-48

Query Match          39.6%; Score 93.5; DB 9; Length 21;
Best Local Similarity 63.0%; Pred. No. 0.00039;
Matches 17; Conservative 0; Mismatches 1; Indels 9; Gaps 1;

QY 1 ILKRPWMPWRRKHEAPEPEPMILKK 27
DB 1 ILKRPWMPWRRKD-----MILKK 18

; RESULT 7
; US-09-030-619-95
; Sequence 95, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erle, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTI-BIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030.619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 95
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cationic Peptide Analogue
; US-09-030-619-95

Query Match          38.6%; Score 91; DB 9; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00048;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 ILKRPWMPWRRK 36
DB 1 ILKRPWMPWRRK 13

; RESULT 8
; US-09-030-619-99
; Sequence 99, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erle, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTI-BIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030.619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 99
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cationic Peptide Analogue
; US-09-030-619-99

Query Match          38.6%; Score 91; DB 9; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00048;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 ILKRPWMPWRRK 36
DB 1 ILKRPWMPWRRK 13

; RESULT 9
; US-10-252-773-3
; Sequence 3, Application US/10252773
; Publication No. US2003013183A1
; GENERAL INFORMATION:
; APPLICANT: EVERETT, NICHOLAS P.
; APPLICANT: LI, JOUNGSHUN
; APPLICANT: LAWRENCE, CHRISTOPHER
; APPLICANT: DAVIES, MAELOR H.
; TITLE OF INVENTION: PEPTIDES WITH ENHANCED STABILITY TO PROTEASE
; TITLE OF INVENTION: DEGRADATION
; FILE REFERENCE: INTERLINK 3.0-003
; CURRENT APPLICATION NUMBER: US/10/252.773
; CURRENT FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 60/106.373
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/106.573
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: antimicrobial peptide
; US-10-252-773-3

Query Match          38.6%; Score 91; DB 16; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00048;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 ILKRPWMPWRRK 36
DB 1 ILKRPWMPWRRK 13

; RESULT 10
; US-09-030-619-72
; Sequence 72, Application US/09030619B
; Patent No. US20020035061A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-72
```

```
Query Match          38.6%; Score 91; DB 9; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00052;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      24 ILKKMPWMPWRRK 36
Db      1 ILKKMPWMPWRRK 13
```

```
RESULT 11
US-09-030-619-108
; Sequence 108, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 108
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-108
```

```
Query Match          38.6%; Score 91; DB 9; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00052;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      24 ILKKMPWMPWRRK 36
Db      1 ILKKMPWMPWRRK 13
```

```
RESULT 12
US-09-030-619-69
; Sequence 69, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 69
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-69
```

```
Query Match          38.6%; Score 91; DB 9; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.00074;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      24 ILKKMPWMPWRRK 36
Db      1 ILKKMPWMPWRRK 13
```

```
RESULT 13
US-09-030-619-67
; Sequence 67, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 67
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-67
```

```
Query Match          36.9%; Score 87; DB 9; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      25 LKKMPWMPWRRK 36
Db      1 LKKMPWMPWRRK 12
```

```
RESULT 14
US-09-030-619-23
; Sequence 23, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
```

APPLICANT: Taylor, Robert
APPLICANT: Erfle, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
TITLE OF INVENTION: WITH ANTIBIOTICS
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030.619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 23
LENGTH: 12
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-23

Query Match 36.4%; Score 86; DB 9; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILKKPMWPMRRK 12
|||||
Db 1 ILKKPMWPMRRK 12

RESULT 15
US-09-030-619-112
Sequence 112; Application US/09030619B
Patent No. US20020035061A1
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfle, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
TITLE OF INVENTION: WITH ANTIBIOTICS
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030.619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 112
LENGTH: 12
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-112

Query Match 36.4%; Score 86; DB 9; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24 ILKKPMWPMRR 35
|||||
Db 1 ILKKPMWPMRR 12

Search completed: July 29, 2003, 16:46:46
Job time : 33.0411 secs

THIS PAGE BLANK (USPTO)

polyoxyalkylene-modified cationic peptides, useful for treating tumours

Claim 1; Page 15; 94pp: English.

This sequence represents a cationic peptide amino acid sequence, which can be used in the pharmaceutical composition of the invention. The invention relates to a pharmaceutical composition containing at least one activated polyoxyalkylene (APO)-modified cationic peptide. The modification of peptides with APO increases their activity against tumour cells, including those with a multidrug resistant phenotype. The pharmaceutical composition can be used to treat tumours, specifically, lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary, cervix, uterus, skin, prostate, liver and colon.

Sequence 28 AA:

Query Match 58.7%; Score 141.5; DB 21; Length 28;
Best Local Similarity 58.8%; Pred. No. 4.2e-11;
Matches 20; Conservative 2; Mismatches 1; Indels 11; Gaps 1;

4 KMPWMPRRKHEAPEPIMLTKMPWMPRRK 37
:|||||||
3 RMPWMPRRK-----ILMRWMPWMPRRK 25

RESULT 2

AAW6363
ID AAW6363 standard; peptide; 27 AA.

AAW6363;

12-JAN-1999 (first entry)

Indolicidin analogue MBI 11B20:

Indolicidin analogue; resistance; cationic peptide; antibiotic;
bacterial infection; tolerance; antibacterial; microorganism;
bacteria; fungus; parasite; virus.

Bos taurus.
OS Synthetic.

WO9840401-A2.

17-SEP-1998.

10-MAR-1998; 98WO-CA00190.

25-FEB-1998; 98US-0030619.

10-MAR-1997; 97US-0040649.

20-AUG-1997; 97US-0915314.

26-SEP-1997; 97US-0060099.

(MICR-) MICROLOGIX BIOTECH INC.

Fraser JR, McNicol PJ, West MHP;

WPI: 1998-520800/44.

New indolicidin peptide analogues - useful for, e.g. enhancing

activity of antibiotic or overcoming tolerance, acquired resistance

or inherent resistance of microorganisms

Claim 1; Page 91; 105pp: English.

The present sequence represents an indolicidin analogue. The present invention describes compositions and methods for treating infection, especially bacterial infections. The compositions and methods use cationic peptides in combination with an antibiotic agent which are then administered to a patient to enhance the activity of the antibiotic agent, to overcome: (a) tolerance; (b) acquired resistance; and (c) inherent resistance. The combinations of antibiotics and cationic

peptides can provide synergistic activity against a microorganism that is tolerant, inherently resistant, or has acquired resistance to an antibiotic agent. They can be used for killing e.g. bacteria, fungi, parasites and viruses.

Sequence 27 AA:

Query Match 57.1%; Score 137.5; DB 19; Length 27;
Best Local Similarity 61.8%; Pred. No. 1.3e-10;
Matches 21; Conservative 2; Mismatches 0; Indels 11; Gaps 2;

4 KMPWMPRRKHEAPEPIMLTKMPWMPRRK 37
:|||||||
3 RMPWMPRRK-----ML-RMPWMPRRK 25

RESULT 3

AAV44668
ID AAV44668 standard; Protein; 63 AA.

AAV44668;

18-APR-2000 (first entry)

Poly-(Indol (1-13)-Met-Ala-Arg-Ile-Ala-Met)3 protein.

Crosslinked indolicidin analog; X-indolicidin; poly-Indol 1-13;
stability; bovine neutrophil; antimicrobial; antibacterial; fungicide;
KW protozoacide; virucide; anti-HIV; human immunodeficiency virus-1;
KW HIV-1; gram positive bacteria; gram negative; Staphylococcus aureus;
KW Escherichia coli; Salmonella typhimurium; yeast; fungi; protozoa;
KW Candida albicans; Cryptococcus neoformans; Giardia; Acanthamoeba;
hexapeptide spacer.

Synthetic.

Bos sp.

Location/Qualifiers

1..5 /label= Enterokinase_recognition_site

5..6 /label= Enterokinase_cleavage_site

6..7 /label= Cyanogen_bromide_cleavage_site

20..21 /label= Cyanogen_bromide_cleavage_site

25..26 /label= Cyanogen_bromide_cleavage_site

39..40 /label= Cyanogen_bromide_cleavage_site

44..45 /label= Cyanogen_bromide_cleavage_site

58..59 /label= Cyanogen_bromide_cleavage_site

20..25 /label= Hexapeptide_spacer

39..44 /label= Hexapeptide_spacer

58..63 /label= Hexapeptide_spacer

WO9965510-A1.

23-DEC-1999.

20-MAY-1999; 99WO-US11165.

18-JUN-1998; 98US-0099631.

(REGC) UNIV CALIFORNIA.

Selsted ME, Osapay K;

```

DR      NP-2000-147133/13.
DR      N-PSDB; AA249764.
XX      Crosslinked indolicidin analogs with antimicrobial activity against
PT      bacteria, yeast, fungi, protozoa and viruses
PS      Example 1C; Fig 1; 53pp; English.
XX

CC      The patent discloses crosslinked analogs of indolicidin (Indol 1-13)
CC      which is a naturally occurring peptide isolated from bovine neutrophils
CC      and has antimicrobial activity. The crosslinked indolicidin
CC      (X-indolicidin) analogs are stable and have antimicrobial activity
CC      against gram positive and negative bacteria (e.g. Staphylococcus aureus,
CC      Escherichia coli and Salmonella typhimurium), yeasts and fungi (e.g.
CC      Candida albicans, Cryptococcus neoformans), protozoa (e.g. Giardia
CC      species and Acanthamoeba species), and viruses (e.g. HIV-1).
CC      They can be used for reducing or inhibiting the growth or survival of
CC      microorganisms in an environment e.g. a food or food product, a
CC      solution, an inanimate object comprising a surface, or a mammal.
CC      The present sequence is a protein comprising three
CC      copies of Indol 1-13 each separated by a hexapeptide spacer sequence.
CC      A recombinant construct encoding this sequence was used for the
CC      expression of Indol-homoserine (Hse) analog. The ability of
CC      Indol-Hse analog to maintain antimicrobial activity provides a means to
CC      produce X-indolicidin analog precursors in sufficient quantities.
XX      SQ      Sequence      63 AA:
XX
XX      Query Match      55.4%; Score 133.5; DB 21; Length 63;
XX      Best Local Similarity 64.7%; Pred. No. 1e-09;
XX      Matches      22; Conservative 0; Mismatches 5; Indels 7; Gaps 2
XX
XX      4 KWPMPWPMRRKHEAPEAEPIIMILK-KWPMPWPMRR 36
XX      |||||
XX      11 KWPMPWPMRRM-----ARIAMILPWPMPWPMRR 38
XX
XX      RESULT 4
XX      ID      AA57142
XX      AAY57142 standard; Protein; 63 AA.
XX
XX      AAY57142:
XX
XX      28-FEB-2000 (first entry)
XX
XX      Indolicidin fusion peptide amino acid sequence.
XX
XX      Indolicidin analogue; antimicrobial activity; helminth; bacteria; virus;
XX      treatment; inhibit growth; micro-organism; contact lens solution;
XX      transgenic plant; surgical instrument; yeast; fungi; protozoa.
XX
XX      Synthetic.
XX
XX      WO9558141-A1.
XX
XX      PN      18-NOV-1999.
XX
XX      PD      05-MAY-1999; 99WO-US09942.
XX
XX      PF      12-MAY-1998; 98US-0076227.
XX
XX      PR      (REGC ) UNITV CALIFORNIA.
XX
XX      PA      Selsted ME.
XX
XX      PI      WPI: 2000-053028/04.
XX      DR      N-PSDB; AA245123.
XX
XX      New indolicidin analogues, active against bacteria, yeast, fungi,
XX      protozoa and virus, used for, e.g. treating infections
XX
XX      Disclosure: Fig 6; 62pp; English.
XX

```

Query Match	Best Local Similarity	55.4%	Score 133.5	DB 21:	Length 63:
Matches	22:	Conservative	0:	Mismatches	5; Indels 7; Gaps 2
OY	4	KWPMPWPKRRKHEAPEEAPRIMILK-KMPMPWPKRR	36		
DB	11	KWPMPWPKRRM-----ARIAMLLPKKMPMPWPKRR	38		
RESULT 5					
AAW12899					
ID	AAW12899	standard; peptide; 16 AA.			
XX	AAW12899;				
AC					
XX					
XX	10-DEC-1997	(first entry)			
DE					
XX	Antimicrobial cationic peptide CP-26.				
XX					
KW	Bacterial; viral; antitumor; food; preservative; inhibitor; growth;				
KW	bacterium; yeast; endotoxaemia; sepsis; antibiotic; fungal;				
KW	antiviral; Candida albicans; steriliant; Salmonella; yersina;				
KW	Shigella.				
XX					
OS	Synthetic.				
XX					
PN	WO9708199-A2.				
PD					
XX	06-MAR-1997.				
PF					
XX	23-AUG-1996;	96WO-IB00996.			
PR					
XX	23-AUG-1995;	95US-0002687.			
XX					
PA	(UYBR-) UNITV BRITISH COLUMBIA.				
XX					
PI	Falla TJ, Gough M, Hancock RM;				

DR WPT; 2000-223549/19.

[illegible]

CC	in (Viral at least 2 X = F or Y. The analogues are used to treat
CC	infections caused by bacteria (Gram positive or negative, or anaerobic);
CC	fungi (yeast or moulds); parasites (protozoa, nematodes, cestodes or
CC	trematodes) or viruses. Typical of very many pathogens that can be
CC	controlled are Leishmania, Trypanosoma, Ascaris lumbricoides, Fasciola
CC	hepatica, Klebsiella pneumoniae, Bordetella pertussis, Staphylococcus
CC	aureus, Listeria, Clostridium, rotavirus and papilloma virus. Compounds
CC	derived from the analogues may be used similarly; the compounds may
CC	also be prepared from antibiotics or antiparasitic agents. The analogues
CC	may be used therapeutically or to coat medical devices; also they are
CC	useful as surface disinfectants, as additives to shampoo or soaps, as
CC	insecticides or herbicides, or as preservatives for foods and technical
CC	materials. The analogues are administered by injection, lavage, orally
CC	or topically, generally at 0.1-50 mg/kg. These analogues have a broader
CC	spectrum of activity than indolicidin and modification as compounds
CC	reduces their toxicity.
CC	
XX	
SO	Sequence 21 AA;
QY	Query Match 40.7%; Score 98; DB 19; Length 21;
Db	Best Local Similarity 64.3%; Pred. No. 9, 2e-06;
	Matches 18; Conservative 0; Mismatches 0; Indels 10; Gaps 1
1	1 ILKKWPMWPMRRRHEAPEPEIMILKK 28
1	1 ILKKWPMWPMRR-----IMILKK 18
AA9Y1808	
AA9Y1808	standard; Peptide: 21 AA.
AA9Y1808;	
06-JUN-2000	(first entry)
Amino acid sequence of cationic peptide MBI 1106CN.	
Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;	
leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma;	
breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;	
multidrug resistance.	
Synthetic.	
WO9965506-A2.	
23-DEC-1999.	
14-JUN-1999;	99WO-CA00552.
12-JUN-1998;	98US-0096541.
(MICR-) MICROLOGIX BIOTECH INC.	
Friedland HD, Krieger TJ, Taylor R, Effle D, Fraser JR, West MHP;	
WPI: 2000-223549/19.	
Novel pharmaceutical composition containing optionally activated	
polyoxyalkylene-modified cationic peptides, useful for treating tumours	
Disclosure: Page 15; 9app; English.	
This sequence represents a cationic peptide amino acid sequence, which	
can be used in the pharmaceutical composition of the invention. The	
invention relates to a pharmaceutical composition containing at least one	
activated polyoxyalkylene (APO)-modified cationic peptide. The	
modification of peptides with APO increases their activity against tumour	
cells, including those with a multidrug resistant phenotype. The	
pharmaceutical composition can be used to treat tumours, specifically	
lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,	

CC cervix, uterus, skin, prostate, liver and colon.
XX
S0 Sequence 21 AA; 40.7%; Score 98; DB 21; Length 21;
Query Match Best Local Similarity 64.3%; Pred. NO. 9.2e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
OY 1 ILKKPWPMPRRKHEAPEAPRIMLKK 28
1 ILKKPWPMPRRKHEAPEAPRIMLKK 18
Db 1 ILKKPWPMPRR-----IMLKK 18
RESULT 10
ID AAY24570
AY24570 standard; peptide; 20 AA.
XX
AC AAY24570;
XX
DT 18-AUG-1999 (first entry)
XX
DE Indolicidin analogue #22.
XX
KW Indolicidin; bacterial infection; photo-oxidised solubiliser;
KW antimicrobial; antibiotic; antiarrhythmic; surface disinfectant;
KW additive; shampoo; soap; insecticide; herbicide; preservative;
KW food; technical material.
XX
OS Synthetic.
XX
MO9807745-A2.
XX
PN 26-FEB-1998.
PD
PE 21-AUG-1997; 97WO-US14779.
PF
PR 13-JAN-1997; 97US-0034949.
PR 21-AUG-1996; 96US-0024754.
XX
PA (MICR-) MICROLOGIX BIOTECH INC.
XX
PI Erfle D, Fraser JR, Krieger TJ, Taylor R, West MH;
XX
DR WPI: 1998-169090/15.
XX
XX New indolicidin analogues with antimicrobial activity and related
XX nucleic acid - vectors, transformed cells and antibodies, also
XX conjugates with polyoxalkylene glycol and fatty acid to reduce
XX toxicity, useful therapeutically, as disinfectants etc.
XX
PS Claim 12; Page 89; 129pp; English.
XX
XX AAY24549 to AAY24615 represent indolicidin analogues of formulae
XX (I)-(VIII) containing up to 25 amino acids (aa): RX2XXXB (I), BX2XXXB
XX (II), BB2XXXB (III), BX2XXBBN(AA)NMLBAGS (IV), BX2XXBB(AA)N
XX (V), LB2XXBBN(AA)NMLBAGS (VI), LK2XXBBN(AA)NMLBAGS (VII),
XX Where 2 = P or V; X = hydrophobic residue, preferably W; B = basic aa,
XX preferably R or K; AA = any aa; n = 0 or 1; in (II), at least 1 Z = V;
XX in (VIII) at least 2 X = F or Y. The analogues are used to treat
XX infections caused by bacteria (gram positive or negative, cestodes or
XX fungi (yeast or moulds); parasites (protozoa, nematodes, cestodes or
XX trematodes) or viruses. Typical of very many pathogens that can be
XX controlled are leishmania, Trypanosoma, Ascaris lumbricoides, Fasciola
XX hepatica, Klebsiella pneumoniae, Bordetella pertussis, Staphylococcus
XX aureus, Listeria, Clostridium, rotavirus and papilloma virus. Compounds
XX derived from the analogues may be used similarly; the compounds may
XX also be prepared from antibiotics or antiarrhythmic agents. The analogues
XX may be used therapeutically, or to coat medical devices; also they are
XX useful as surface disinfectants, as additives to shampoo or soaps, as
XX insecticides or herbicides, or as preservatives for foods and technical
XX materials. The analogues are administered by injection, lavage, orally
XX or topically, generally at 0.1-50 mg/kg. These analogues have a broader
XX spectrum of activity than indolicidin and modification as compounds

CC		reduces their toxicity.
XX	Sequence	20 AA;
SO		
	Query Match	38.8%; Score 93.5; DB 19; Length 20;
	Best Local Similarity	60.7%; Pred No. 3.2e-05;
	Matches 17; Conservative	0; Mismatches 0; Indels 11; Gaps 1;
OY		
	1 ILKKPMPWPKRRKHEAPEAPETMLKK 28	
	1 ILKKMPWMPWR-----MILKK 17	
DB		
	RESULT 11	
	AA91807	
ID	AA91807 standard; Peptide; 20 AA.	
XX		
AC	AA91807;	
XX		
DT	06-JUN-2000 (first entry)	
XX		
DE	Amino acid sequence of cationic peptide MBI 11DSCN.	
XX		
KM	Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;	
KM	leukemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;	
KM	breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;	
KW	multidrug resistance.	
OS	Synthetic.	
XX		
PN	WO965506-A2.	
XX		
PD	23-DEC-1999.	
XX		
PE	14-JUN-1999; 99WO-CA00552.	
XX		
PR	12-JUN-1998; 98US-0096541.	
XX		
PA	(MICR-) MICROLOGIX BIOTECH INC.	
XX		
PI	Friedland HD, Krieger TJ, Taylor R, Erfile D, Fraser JR, West MHF;	
XX		
DR	WPI; 2000-223549/19.	
XX		
Pt	Novel pharmaceutical composition containing optionally activated	
Pt	polyoxalkylene-modified cationic peptides, useful for treating tumours	
-		
PS	Disclosure; Page 15; 94pp; English.	
XX		
CC	This sequence represents a cationic peptide amino acid sequence, which	
CC	can be used in the pharmaceutical composition of the invention. The	
CC	invention relates to a pharmaceutical composition containing at least one	
CC	activated polyoxalkylene (APO)-modified cationic peptide. The	
CC	modification of peptides with APO increases their activity against tumour	
CC	cells, including those with a multidrug resistant phenotype. The	
CC	pharmaceutical composition can be used to treat tumours, specifically	
CC	lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,	
CC	cervix, uterus, skin, prostate, liver and colon.	
XX		
SO	Sequence	20 AA;
	Query Match	38.8%; Score 93.5; DB 21; Length 20;
	Best Local Similarity	60.7%; Pred. No. 3.2e-05;
	Matches 17; Conservative	0; Mismatches 0; Indels 11; Gaps 1;
OY		
	1 ILKKPMPWPKRRKHEAPEAPETMLKK 28	
	1 ILKKMPWMPWR-----MILKK 17	
DB		
	RESULT 12	
	AA91807	
ID	AA91807 standard; Peptide; 20 AA.	
XX		
AC	AA91807;	
XX		
DT	06-JUN-2000 (first entry)	
XX		
DE	Amino acid sequence of cationic peptide MBI 11DSCN.	
XX		
KM	Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;	
KM	leukemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;	
KM	breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;	
KW	multidrug resistance.	
OS	Synthetic.	
XX		
PN	WO965506-A2.	
XX		
PD	23-DEC-1999.	
XX		
PE	14-JUN-1999; 99WO-CA00552.	
XX		
PR	12-JUN-1998; 98US-0096541.	
XX		
PA	(MICR-) MICROLOGIX BIOTECH INC.	
XX		
PI	Friedland HD, Krieger TJ, Taylor R, Erfile D, Fraser JR, West MHF;	
XX		
DR	WPI; 2000-223549/19.	
XX		
Pt	Novel pharmaceutical composition containing optionally activated	
Pt	polyoxalkylene-modified cationic peptides, useful for treating tumours	
-		
PS	Disclosure; Page 15; 94pp; English.	
XX		
CC	This sequence represents a cationic peptide amino acid sequence, which	
CC	can be used in the pharmaceutical composition of the invention. The	
CC	invention relates to a pharmaceutical composition containing at least one	
CC	activated polyoxalkylene (APO)-modified cationic peptide. The	
CC	modification of peptides with APO increases their activity against tumour	
CC	cells, including those with a multidrug resistant phenotype. The	
CC	pharmaceutical composition can be used to treat tumours, specifically	
CC	lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,	
CC	cervix, uterus, skin, prostate, liver and colon.	
XX		
SO	Sequence	20 AA;
	Query Match	38.8%; Score 93.5; DB 21; Length 20;
	Best Local Similarity	60.7%; Pred. No. 3.2e-05;
	Matches 17; Conservative	0; Mismatches 0; Indels 11; Gaps 1;
OY		
	1 ILKKPMPWPKRRKHEAPEAPETMLKK 28	
	1 ILKKMPWMPWR-----MILKK 17	
DB		
	RESULT 12	
	AA91807	
ID	AA91807 standard; Peptide; 20 AA.	
XX		
AC	AA91807;	
XX		
DT	06-JUN-2000 (first entry)	
XX		
DE	Amino acid sequence of cationic peptide MBI 11DSCN.	
XX		
KM	Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;	
KM	leukemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;	
KM	breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;	
KW	multidrug resistance.	
OS	Synthetic.	
XX		
PN	WO965506-A2.	
XX		
PD	23-DEC-1999.	
XX		
PE	14-JUN-1999; 99WO-CA00552.	
XX		
PR	12-JUN-1998; 98US-0096541.	
XX		
PA		

```

ID      AA12873 standard; peptide: 13 AA.
XX
AC      AAM12873;
XX
XX      10-DEC-1997 (first entry)
XX
DE      Antimicrobial cationic peptide CP-11.
XX
XX      Bacterial; viral; antitumour; food; preservative; inhibitor; growth;
XX      bacterium; yeast; endotoxaemia; sepsis; antibiotic; fungal;
XX      antiviral; Candida albicans; steriliant; Salmonella; Yersina;
XX      Shigella.
XX
OS      Synthetic.
XX
XX      W09708199-A2.
XX
XX      06-MAR-1997.
XX
XX      23-AUG-1996; 96WO-IB00996.
XX
XX      23-AUG-1995; 95US-0002687.
XX
XX      (UYBR-) UNIV BRITISH COLUMBIA.
XX
XX      Falla TJ, Gough M, Hancock RW;
XX
XX      WPI: 1997-179179/16.
XX
XX      Cationic peptide(s) having anti-microbial activity - used for the
XX      inhibition of bacterial and viral growth, as an antitumour agent,
XX      and as a food preservative
XX
XX      Claim 2; Page 65; 89pp; English.
XX
XX
XX      The present sequence represents a specifically claimed novel isolated
XX      cationic peptide which has antimicrobial activity. The amino acid
XX      sequence of antimicrobial cationic peptides (including the present
XX      sequence) is selected from: X1X1ProX2X3X2Pro(X2X2Pro)nX2X3(X5)O;
XX      X1X1ProX2X3X4(X5)ProX2X3X3; X1X1X3(ProTrp)uX3X2X5X2X2X5X2(X5)O;
XX      X1X1X3X3X2Pro(X2X2Pro)nX2(X5)m; where m = 1-5; n = 1-2; o = 2-5; r
XX      = 0-8; u = 0-1; X1 = Ile, Leu, Val, Phe, Tyr, Trp or Met; X2 = Trp or
XX      Phe; X3 = Arg or Lys; X4 = Trp or Lys; and X5 = Phe, Trp, Arg, Lys or
XX      Pro. The peptides are preferably amidated or carboxymethylated. The
XX      peptides may be used in methods for inhibiting the growth of a bacterium
XX      or yeast, or for inhibiting an endotoxaemia or sepsis associated
XX      disorder in a subject. The peptides have a broad activity against
XX      antibiotic resistant bacteria, combined with activity against the
XX      medically important fungus Candida albicans. In addition, the peptides
XX      are useful as antitumour agents and/or antiviral agents. The peptides
XX      may be used as sterilants or preservatives of materials susceptible to
XX      microbial or viral contamination, e.g. in processed foods to inhibit
XX      Salmonella, Yersina and Shigella. The peptides are compact and tend to
XX      have a unique polypoline type II extended helix structure that permits
XX      them to span the membrane with relatively few amino acids. The peptides
XX      possess the ability to work synergistically with antibiotics, and in
XX      addition, some of them possess anti-endotoxin activity.
XX      N.B. The present sequence represents SEQ ID NO:1 in the claims and
XX      examples of the specification, but differs slightly from the SEQ ID NO:1
XX      in the sequence listing on page 51 of the specification (see AAM27179).
XX
XX      Sequence 13 AA:
XX
XX      Query Match 37.8%; Score 91; DB 18; Length 13;
XX      Best Local Similarity 100.0%; Pred. NO. 4,1e-05;
XX      Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      1 ILKKPMPMPWRRK 13
XX      |||||
XX      1 ILKKPMPMPWRRK 13

```

```

ID      AAY24609
XX      AAY24609 standard; peptide: 13 AA.
AC      AAY24609:
XX
DT      18-AUG-1999 (first entry)
XX
DE      Indolicidin analogue #61.
XX
KW      Indolicidin; bacterial infection; photo-oxidised solubiliser;
KW      antimicrobial; antibiotic; antiarrhythmic; surface disinfectant;
KW      additive; shampoo; soap; insecticide; herbicide; preservative;
KW      food; technical material.
XX
OS      Synthetic.
XX
PN      WC9807745-A2.
XX
PD      26-FEB-1998.
XX
PF      21-AUG-1997: 97WO-US14779.
XX
PR      13-JAN-1997: 97US-0034949.
PR      21-AUG-1996: 96US-0024754.
XX
PA      (MICR-) MICROLOGIX BIOTECH INC.
XX
PI      Erfle D, Fraser JR, Krieger TJ, Taylor R, West MH:
XX      WPI; 1998-169090/15.
XX
PT      New indolicidin analogues with antimicrobial activity and related
PT      nucleic acid - vectors, transformed cells and antibodies, also
PT      conjugates with polyoxyalkylene glycol and fatty acid to reduce
PT      toxicity, useful therapeutically, as disinfectants etc.
XX
PS      Example 1: Page 32; 129pp: English.
XX
AA      AAY24549 to AAY24615 represent indolicidin analogues of formulae
CC      (i)-(viii) containing up to 25 amino acids (aa): RXZXXZXB (I), BXZXXZXB
CC      (II), BBHXXZXXZXB (IIi), BXZXXZXBBA(AA)nmLBBAGS (IV), BXZXXZXBBA(AA)nm
CC      (V), LBBHXXZXXZXXNRK (VI), LKHXZXXZXXNRK (VII) and BBXZXXZXBBA (VIII).
CC      Where Z = P or V; X = hydrophobic residue, preferably W; B = basic aa,
CC      preferably R or K; AA = any aa; n = 0 or 1; in (II), at least 1 Z = V;
CC      in (VIII) at least 2 X = F or Y. The analogues are used to treat
CC      infections caused by bacteria (Gram positive or negative, or anaerobic);
CC      fungi (yeast or moulds); parasites (protozoa, nematodes, cestodes or
CC      trematodes) or viruses. Typical of very many pathogens that can be
CC      controlled are Leishmania, Trypanosoma, Ascaris lumbricoides, Fasciola
CC      hepatica, Klebsiella pneumoniae, Bordetella pertussis, Staphylococcus
CC      aureus, listeria, Clostridium, rotavirus and papilloma virus. Compounds
CC      derived from the analogues may be used similarly: the compounds may
CC      also be prepared from antibiotics or antiarrhythmic agents. The analogues
CC      may be used therapeutically or to coat medical devices; also they are
CC      useful as surface disinfectants, as additives to shampoo or soaps, as
CC      insecticides or herbicides, or as preservatives for foods and technical
CC      materials. The analogues are administered by injection, lavage, orally
CC      or topically, generally at 0.1-50 mg/kg. These analogues have a broader
CC      spectrum of activity than indolicidin and modification as compounds
CC      reduces their toxicity.
XX
SQ      Sequence 13 AA:
XX
Query Match 37.8%; Score 91; DB 19; Length 13;
Best Local Similarity 100.0%; Prod. No. 4.1e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 ILKKPMPMPWRRK 13
        |||||
DB      1 ILKKPMPMPWRRK 13

```

```

AAW6378
ID AAW6378 standard; peptide: 13 AA.
XX
AC AAW6378;
XX
DT 12-JAN-1999 (first entry)
XX
DE Cationic peptide of claim 15 #5.
XX
XX Indolicidin analogue; resistance: cationic peptide; antibiotic;
KM bacterial infection; tolerance; antibacterial; microorganism;
KM bacteria; fungus; parasite; virus.
XX
OS Synthetic.
XX
PN W09840401-A2.
XX
PD 17-SEP-1998.
XX
PF 10-MAR-1998; 98WO-CA00190.
XX
PR 25-FEB-1998; 98US-0030619.
PR 10-MAR-1997; 97US-0040649.
PR 20-AUG-1997; 97US-0915314.
PR 26-SEP-1997; 97US-0060099.
XX
PA (MTCR-) MICROLOGIX BIOTECH INC.
XX
PI Fraser JR, McNicol PJ, West MHP;
XX
DR WPI: 1998-520800/44.
XX
PT New indolicidin peptide analogues - useful for, e.g. enhancing
PT activity of antibiotic or overcoming tolerance, acquired resistance
PT or inherent resistance of microorganisms
XX
XX
SQ Claim 15; Page 93; 105pp; English.
XX
CC The present sequence represents a specifically claimed cationic peptide
CC from the present invention. The present invention describes compositions
CC and methods for treating infection, especially bacterial infections. The
CC compositions and methods use cationic peptides in combination with an
CC antibiotic agent which are then administered to a patient to enhance the
CC activity of the antibiotic agent, to overcome: (a) tolerance; (b)
CC acquired resistance; and (c) inherent resistance. The combinations of
CC antibiotics and cationic peptides can provide synergistic activity
CC against a microorganism that is tolerant, inherently resistant, or has
CC acquired resistance to an antibiotic agent. They can be used for killing
CC e.g. bacteria, fungi, parasites and viruses.
XX
SQ Sequence 13 AA;
XX
Query Match 37.8%; Score 91; DB 19; Length 13;
Best Local Similarity 100.0%; Pred. NO. 4.1e-05; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKPMPMPRRK 13
   | | | | | | | | | | | | | |
Db 1 ILKKPMPMPRRK 13

RESULT 15
AAW71690
ID AAW71690 standard; Peptide: 13 AA.
XX
AC AAW71690;
XX
DT 11-JAN-1999 (first entry)
XX
DE Cationic peptide MB11 (MW 1879).
XX
XX MB11; cationic peptide; plasmid pK1; small cryptic plasmid;
KM replication; RepA; vector; RAMP.

```

```

XX
OS Synthetic.
XX
PN W09841636-A2.
XX
PD 24-SEP-1998.
XX
PF 16-MAR-1998; 98WO-CA00214.
XX
PR 14-MAR-1997; 97US-0040722.
XX
PA (BUR1/) BURIAN J.
PA (KAYW/) KAY W W.
XX
PI Burian J, Kay WW;
XX
DR WPI: 1998-531571/45.
XX
PT Increasing plasmid copy number in a cell with the repA gene product
PT - and a small cryptic plasmid ori sequence, useful for high level
PT expression of e.g. cytokines, antigens or therapeutic proteins
XX
XX
PS Example 13; Page 54; 82pp; English.
XX
CC MB11 is a small (mol.wt. 1879) cationic peptide. DNA encoding
CC MB11 has been incorporated into vector pR2h-B1, in which the
CC replication leader (R21) sequence of RepA (see also AAW71686) is
CC joined to 2 Hpro peptides (see also AAW71692), to provide a
CC vector for expression of MB11 in host cells. The invention
CC provides controlled replication plasmid vectors (RAMP vectors)
CC comprising a replication origin of a small cryptic plasmid and a
CC gene encoding RepA. The vectors can reach very high levels of
CC plasmid replication, but are not lethal to the host cell, and can
CC be used to direct the high level expression of e.g. cytokines,
CC antigens and therapeutic proteins.
XX
SQ Sequence 13 AA;
XX
Query Match 37.8%; Score 91; DB 19; Length 13;
Best Local Similarity 100.0%; Pred. NO. 4.1e-05; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKPMPMPRRK 13
   | | | | | | | | | | | | | |
Db 1 ILKKPMPMPRRK 13

Search completed: July 29, 2003, 16:31:40
Job time : 60.3151 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2003, 16:34:31 ; Search time 33.9589 Seconds

(Without alignments)
129.395 Million cell updates/sec

Title: SEQ35-27-35

Perfect score: 241

Sequence: 1 ILKKMPMPWRKHEAPEAPILKKMPMPWRK 37

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA:*
1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppaa/US10C_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	141.5	58.7	28	9	US-09-030-619-104
2	137.5	57.1	28	9	US-09-030-619-50
3	99	41.1	21	9	US-09-030-619-69
4	98	40.7	21	9	US-09-030-619-52
5	93.5	38.8	20	9	US-09-030-619-51
6	91	37.8	13	9	US-09-030-619-95
7	91	37.8	13	9	US-09-030-619-99
8	91	37.8	13	16	US-10-252-773-3
9	91	37.8	14	9	US-09-030-619-72
10	91	37.8	14	9	US-09-030-619-108
11	87.5	36.3	21	9	US-09-030-619-47
12	87	36.1	12	9	US-09-030-619-67
13	86	35.7	12	9	US-09-030-619-112
14	86	35.7	13	9	US-09-030-619-53
15	86	35.7	13	9	US-09-030-619-107

16	86	35.7	13	9	US-09-030-619-109	Sequence 109, App
17	86	35.7	14	9	US-09-030-619-54	Sequence 54, App
18	86	35.7	14	9	US-09-030-619-110	Sequence 110, App
19	85	35.3	12	9	US-09-030-619-73	Sequence 73, App
20	85	35.3	13	9	US-09-030-619-41	Sequence 41, App
21	85	35.3	13	9	US-09-030-619-103	Sequence 103, App
22	83.5	34.6	15	9	US-09-030-619-39	Sequence 39, App
23	83	34.4	12	9	US-09-030-619-44	Sequence 44, App
24	83	34.4	13	9	US-09-030-619-105	Sequence 105, App
25	83	34.4	13	9	US-09-030-619-106	Sequence 106, App
26	83	34.4	20	9	US-09-030-619-24	Sequence 24, App
27	82.5	34.2	21	9	US-09-030-619-48	Sequence 48, App
28	82	34.0	13	9	US-09-030-619-58	Sequence 58, App
29	82	34.0	13	9	US-09-030-619-59	Sequence 59, App
30	82	34.0	13	9	US-09-030-619-94	Sequence 94, App
31	82	34.0	13	9	US-09-030-619-102	Sequence 102, App
32	81	33.6	11	9	US-09-030-619-79	Sequence 79, App
33	81	33.6	11	9	US-09-030-619-114	Sequence 114, App
34	78	32.4	11	9	US-09-030-619-78	Sequence 78, App
35	78	32.4	11	9	US-09-030-619-113	Sequence 113, App
36	78	32.4	12	9	US-09-030-619-43	Sequence 43, App
37	77.5	32.2	12	9	US-09-030-619-40	Sequence 40, App
38	77	32.0	12	9	US-09-030-619-42	Sequence 42, App
39	76	31.5	12	9	US-09-030-619-82	Sequence 82, App
40	75	31.1	12	9	US-09-030-619-23	Sequence 23, App
41	75	31.1	12	9	US-09-030-619-30	Sequence 30, App
42	75	31.1	12	9	US-09-030-619-111	Sequence 111, App
43	73.5	30.5	12	9	US-09-030-619-75	Sequence 75, App
44	73.5	30.5	12	9	US-09-030-619-76	Sequence 76, App
45	73	30.3	9	9	US-09-030-619-80	Sequence 80, App

ALIGNMENTS

RESULT 1
US-09-030-619-104
Sequence 104, Application US/09030619B
Patent No. US20020035061A1
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Eraser, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: MCNICOL, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
FILE REFERENCE: 660081, 406
CURRENT APPLICATION NUMBER: US/09/030,619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 104
LENGTH: 28
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-104

Query Match 58.7%; Score 141.5; DB 9; Length 28;
Best Local Similarity 58.8%; Pred. No. 2.8e-09;
Matches 20; Conservative 2; Mismatches 1; Indels 11; Gaps 1;

QY 4 KMPMPWRKHEAPEAPILKKMPMPWRK 37
DB 3 KMPMPWRK-----ILKKMPMPWRK 25
RESULT 2
US-09-030-619-50

```
; Sequence 50, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 50
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-50
```

```
Query Match          57.1%; Score 137.5; DB 9; Length 28;
Best Local Similarity 61.8%; Pred. No. 7.6e-09;
Matches 21; Conservative 2; Mismatches 0; Indels 11; Gaps 2;
```

```
QY      4 KMPMPWRRKHEAPEAEPIMLKKMPWMPWRRK 37
      :|||||||:|||||||:|||||||:
-DB     3 KMPMPWRRK-----MIL-KMPMPWRRK 25
```

```
RESULT 3
US-09-030-619-69
; Sequence 69, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 69
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-69
```

```
Query Match          41.1%; Score 99; DB 9; Length 21;
Best Local Similarity 64.3%; Pred. No. 0.0001;
Matches 18; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
```

```
QY      1 ILKKMPWMPWRRKHEAPEAEPIMLKK 28
      :|||||||:|||||||:|||||||:
DB     1 ILKKMPWMPWRRK-----MILKK 18
```

```
RESULT 4
US-09-030-619-52
; Sequence 52, Application US/09030619B
```

```
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-52
```

```
Query Match          40.7%; Score 98; DB 9; Length 21;
Best Local Similarity 64.3%; Pred. No. 0.00013;
Matches 18; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
```

```
QY      1 ILKKMPWMPWRRKHEAPEAEPIMLKK 28
      :|||||||:|||||||:|||||||:
DB     1 ILKKMPWMPWRR-----MILKK 18
```

```
RESULT 5
US-09-030-619-51
; Sequence 51, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 51
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-51
```

```
Query Match          38.8%; Score 93.5; DB 9; Length 20;
Best Local Similarity 60.7%; Pred. No. 0.00039;
Matches 17; Conservative 0; Mismatches 0; Indels 11; Gaps 1;
```

```
QY      1 ILKKMPWMPWRRKHEAPEAEPIMLKK 28
      :|||||||:|||||||:|||||||:
DB     1 ILKKMPWMPWRR-----MILKK 17
```

```
RESULT 6
US-09-030-619-95
; Sequence 95, Application US/09030619B
; Patent No. US20020035061A1
```

GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfle, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
TITLE OF INVENTION: WITH ANTIBIOTICS
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030.619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 95
LENGTH: 13
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-95

Query Match 37.8%; Score 91; DB 9; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0005;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILKKPMPWMPWRRK 13
Db 1 ILKKPMPWMPWRRK 13

RESULT 7
US-09-030-619-99
Sequence 99, Application US/09030619B
Patent No. US20020035061A1
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfle, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
TITLE OF INVENTION: WITH ANTIBIOTICS
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030.619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 99
LENGTH: 13
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-99

Query Match 37.8%; Score 91; DB 9; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0005;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILKKPMPWMPWRRK 13
Db 1 ILKKPMPWMPWRRK 13

RESULT 8
US-10-252-773-3
Sequence 3, Application US/10252773
Publication No. US2003013183A1
GENERAL INFORMATION:

APPLICANT: EVERETT, NICHOLAS P.
APPLICANT: LI, QUNIGSHUN
APPLICANT: LAWRENCE, CHRISTOPHER
APPLICANT: DAVIES, MAELOR H.
TITLE OF INVENTION: PEPTIDES WITH ENHANCED STABILITY TO PROTEASE
TITLE OF INVENTION: DEGRADATION
FILE REFERENCE: INTERLINK 3.0-003
CURRENT APPLICATION NUMBER: US/10/252.773
CURRENT FILING DATE: 2002-09-23
PRIOR APPLICATION NUMBER: 60/106.373
PRIOR FILING DATE: 1998-10-30
PRIOR APPLICATION NUMBER: 60/106.573
PRIOR FILING DATE: 1998-11-02
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 13
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-252-773-3

Query Match 37.8%; Score 91; DB 16; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0005;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILKKPMPWMPWRRK 13
Db 1 ILKKPMPWMPWRRK 13

RESULT 9
US-09-030-619-72
Sequence 72, Application US/09030619B
Patent No. US20020035061A1
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfle, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
TITLE OF INVENTION: WITH ANTIBIOTICS
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030.619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 72
LENGTH: 14
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-72

Query Match 37.8%; Score 91; DB 9; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00054;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILKKPMPWMPWRRK 13
Db 1 ILKKPMPWMPWRRK 13

RESULT 10
US-09-030-619-108
Sequence 108, Application US/09030619B
Patent No. US20020035061A1

```

GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Ertle, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNICOL, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INJECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
TITLE OF INVENTION: WITH ANTIBIOTICS
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 108
LENGTH: 14
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-108

Query Match          37.8%; Score 91; DB %; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00054;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ILKKPMPWRRK 13
        | | | | | | | | | |
DB      1 ILKKPMPWRRK 13

RESULT 11
US-09-030-619-47
; Sequence 47, Application US/09030619B
; Patent No. US20020035061A1
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Ertle, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNICOL, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INJECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
TITLE OF INVENTION: WITH ANTIBIOTICS
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 47
LENGTH: 21
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-47

Query Match          36.3%; Score 87.5; DB %; Length 21;
Best Local Similarity 60.0%; Pred. No. 0.00197;
Matches 15; Conservative 1; Mismatches 0; Indels 9; Gaps 1;

QY      4 KMPMPWRRKHEAPEAPMILKK 28
        : | | | | | | | | | |
DB      3 RMPMPWRRK-----IMILKK 18

RESULT 12
US-09-030-619-67
; Sequence 67, Application US/09030619B
; Patent No. US20020035061A1
GENERAL INFORMATION:

```

```

APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfile, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
TITLE OF INVENTION: WITH ANTIBIOTICS
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 67
LENGTH: 12
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-67

Query Match
Best Local Similarity 100.0%; Score 87; DB 9; length 12;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 LKKMPMPWRRK 13
1 LKKMPMPWRRK 12
Db

RESULT 13
US-09-030-619-112
; Sequence 112, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 112
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cationic Peptide Analogue
;
US-09-030-619-112

Query Match
Best Local Similarity 100.0%; Score 86; DB 9; length 12;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ILKKMPMPWRR 12
1 ILKKMPMPWRR 12
Db

RESULT 14
US-09-030-619-53
; Sequence 53, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.

```

APPLICANT: Taylor, Robert
APPLICANT: Erfile, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
TITLE OF INVENTION: WITH ANTIBIOTICS
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 53
LENGTH: 13
TYPE: PRP
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-53

Query Match 35.7%: Score 86; DB 9; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKPMPMPWRR 12
|||
1 ILKKPMPMPWRR 12

.

RESULT 15
US-09-030-619-107
Sequence 107, Application US/09030619B
Patent No. US20020035061A1

GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfile, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
TITLE OF INVENTION: WITH ANTIBIOTICS
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 107
LENGTH: 13
TYPE: PRP
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-107

Query Match 35.7%: Score 86; DB 9; Length 13;
Best Local Similarity 41.9%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 18; Gaps 1;

QY 7 WMPWRKHAEPEAPIMILKKPMPMPWRR 37
|||
1 WW-----KKMPMPMPWRR 13

Search completed: July 29, 2003, 16:46:46
Job time : 34.9589 secs

THIS PAGE BLANK (USPTO)